Solutions Manual for A. Zee,
Quantum Field Theory in a Nutshell, 2nd Ed.

Yoni BenTov

Contents

Preface

These are the solutions to the problems in *Quantum Field Theory in a Nutshell*, 2nd Ed., that are not already solved in the back of the book. For most of the problems I provide a detailed solution, while for others I sketch the solution and provide a reference to the literature for further details. Some problems are intentionally open-ended, serving as more of an introduction to the literature rather than as a homework assignment. The goal of Zee's text is not only to teach quantum field theory but also to facilitate the transition from student to researcher.

I thank my colleagues and teachers for helpful discussions in preparing these solutions. In particular, I have benefitted from talking to Jen Cano, Gavin Hartnett, Kurt Hinterbichler, Josh Ilany, Yonah Lemonik, Eugeniu Plamadeala, Yinbo Shi, Joe Swearngin, Benson Way and Chiu-Tien Yu. I am grateful to Joshua Feinberg for providing some of the solutions. I am also indebted to the faculty at the University of California at Santa Barbara for their patience in answering my questions. In particular, I thank David Berenstein, Andreas Ludwig, Ben Monreal, Joe Polchinski, Mark Srednicki, and my thesis adviser, Tony Zee.

Finally, I thank my editor Ingrid Gnerlich at Princeton University Press for providing the opportunity to undertake this project. I also thank my friends and family for their support during the completion of this work.

Yoni BenTov
Apr. 9, 2012

v

I Motivation and Foundation

I.2 Path Integral Formulation

1. Verify (5)

$$\langle q_F | e^{-iHT} | q_I \rangle = \int Dq \, e^{i \int_0^T dt [\frac{1}{2} m \dot{q}^2 - V(q)]} \qquad (5)$$

Solution:

Start with the Hamiltonian $\hat{H} = \frac{1}{2m} \hat{p}^2 + V(\hat{q})$. Let $T = N\epsilon$, with $N \to \infty, \epsilon \to 0, T$ fixed. In this way, split up the time evolution operator into N pieces:

$$e^{-i\hat{H}T} = \underbrace{e^{-i\hat{H}\epsilon} e^{-i\hat{H}\epsilon} \dots e^{-i\hat{H}\epsilon}}_{N \text{ copies}}$$

Use this decomposition in the transition amplitude $\langle q_F | e^{-i\hat{H}T} | q_0 \rangle$, and insert one copy of the identity matrix between each pair of $e^{-i\hat{H}\epsilon}$s:

$$\langle q_F | e^{-iHT} | q_0 \rangle = \langle q_F | e^{-i\hat{H}\epsilon} e^{-i\hat{H}\epsilon} \dots e^{-i\hat{H}\epsilon} | q_0 \rangle$$
$$= \langle q_F | e^{-i\hat{H}\epsilon} 1^{(N-1)} e^{-i\hat{H}\epsilon} 1^{(N-2)} e^{-i\hat{H}\epsilon} \dots e^{-i\hat{H}\epsilon} 1^{(1)} e^{-i\hat{H}\epsilon} | q_0 \rangle$$

The superscript is just a label to keep track of the fact that we have inserted $N-1$ identity matrices. It is convenient to resolve each identity matrix in a complete set of position eigenstates:

$$1^{(i)} = \int_{-\infty}^{\infty} dq_i |q_i\rangle \langle q_i|$$

To see why, consider the matrix element of $e^{-i\hat{H}\epsilon} = 1 - i \left(\frac{1}{2m} \hat{p}^2 + V(\hat{q}) \right) \epsilon + O(\epsilon^2)$ between two position eigenstates, $|q_i\rangle$ and $|q_j\rangle$:

$$\langle q_i | e^{-i\hat{H}\epsilon} | q_j \rangle = \langle q_i | 1 - i \left(\frac{1}{2m} \hat{p}^2 + V(\hat{q}) \right) \epsilon + O(\epsilon^2) | q_j \rangle$$
$$= \langle q_i | 1 - i \left(\frac{1}{2m} \hat{p}^2 + V(q_j) \right) \epsilon + O(\epsilon^2) | q_j \rangle$$
$$= \langle q_i | e^{-i\left(\frac{1}{2m} \hat{p}^2 + V(q_j) \right)\epsilon} | q_j \rangle$$
$$= e^{-iV(q_j)\epsilon} \langle q_i | e^{-i \frac{1}{2m} \hat{p}^2 \epsilon} | q_j \rangle$$

The re-exponentiation is justified by the Baker-Campbell-Hausdorff formula:

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]+(\text{higher commutators})}$$

Set $A = -i\epsilon \frac{\hat{p}^2}{2m}$ and $B = -i\epsilon V(\hat{x})$ in the above to get:

$$e^{-i\epsilon \frac{\hat{p}^2}{2m}} e^{-i\epsilon V(\hat{x})} = e^{-i\epsilon \frac{\hat{p}^2}{2m} - i\epsilon V(\hat{x}) + O(\epsilon^2)}$$

So to $O(\epsilon^2) \to 0$ we can perform the above manipulations.

Now insert an identity matrix to the right of $e^{-i\frac{1}{2m}\hat{p}^2\epsilon}$, but this time resolve it in a complete set of momentum eigenstates:

$$1 = \int_{-\infty}^{\infty} \frac{dp}{2\pi} |p\rangle\langle p|$$

This gives, for the above matrix element, the following:

$$\langle q_i|e^{-i\hat{H}\epsilon}|q_j\rangle = e^{-iV(q_j)\epsilon}\langle q_i|e^{-i\frac{1}{2m}\hat{p}^2\epsilon}\,1\,|q_j\rangle$$

$$= e^{-iV(q_j)\epsilon}\langle q_i|e^{-i\frac{1}{2m}\hat{p}^2\epsilon}\left(\int_{-\infty}^{\infty}\frac{dp}{2\pi}|p\rangle\langle p|\right)|q_j\rangle$$

$$= e^{-iV(q_j)\epsilon}\int_{-\infty}^{\infty}\frac{dp}{2\pi}\langle q_i|e^{-i\frac{1}{2m}\hat{p}^2\epsilon}|p\rangle\langle p|q_j\rangle$$

$$= e^{-iV(q_j)\epsilon}\int_{-\infty}^{\infty}\frac{dp}{2\pi}e^{-i\frac{1}{2m}p^2\epsilon}\underbrace{\langle q_i|p\rangle}_{e^{iq_ip}}\underbrace{\langle p|q_j\rangle}_{e^{-iq_jp}}$$

$$= e^{-iV(q_j)\epsilon}\int_{-\infty}^{\infty}\frac{dp}{2\pi}e^{-i\frac{1}{2m}p^2\epsilon+ip(q_i-q_j)}$$

Now do all of this for the original transition amplitude. This will require $N-1$ resolutions of the identity in position eigenstates, as already indicated, and it will require N resolutions of the identity in momentum eigenstates.

$$\langle q_F|e^{-i\hat{H}T}|q_0\rangle =$$
$$\langle q_F|e^{-i\hat{H}\epsilon}\,1^{(N-1)}\,e^{-i\hat{H}\epsilon}\,1^{(N-2)}\,...\,1^{(1)}\,e^{-i\hat{H}\epsilon}\,|q_0\rangle =$$
$$\langle q_F|e^{-i\hat{H}\epsilon}\left[\int_{-\infty}^{\infty}dq_{N-1}|q_{N-1}\rangle\langle q_{N-1}|\right]e^{-i\hat{H}\epsilon}\left[\int_{-\infty}^{\infty}dq_{N-2}|q_{N-2}\rangle\langle q_{N-2}|\right]...\left[\int_{-\infty}^{\infty}dq_1|q_1\rangle\langle q_1|\right]e^{-i\hat{H}\epsilon}|q_0\rangle$$
$$= \int_{-\infty}^{\infty}dq_{N-1}...dq_1\langle q_F|e^{-i\hat{H}\epsilon}|q_{N-1}\rangle\langle q_{N-1}|e^{-i\hat{H}\epsilon}|q_{N-2}\rangle\langle q_{N-2}|\,...\,|q_2\rangle\langle q_2|e^{-i\hat{H}\epsilon}|q_1\rangle\langle q_1|e^{-i\hat{H}\epsilon}|q_0\rangle$$

- $\langle q_F | e^{-i\hat{H}\epsilon} | q_{N-1} \rangle = e^{-iV(q_{N-1})\epsilon} \int_{-\infty}^{\infty} \frac{dp_{N-1}}{2\pi} e^{-i\frac{p_{N-1}^2}{2m}\epsilon + ip_{N-1}(q_F - q_{N-1})}$

- $\langle q_{N-1} | e^{-i\hat{H}\epsilon} | q_{N-2} \rangle = e^{-iV(q_{N-2})\epsilon} \int_{-\infty}^{\infty} \frac{dp_{N-2}}{2\pi} e^{-i\frac{p_{N-2}^2}{2m}\epsilon + ip_{N-2}(q_{N-1} - q_{N-2})}$

    ...

- $\langle q_2 | e^{-i\hat{H}\epsilon} | q_1 \rangle = e^{-iV(q_1)\epsilon} \int_{-\infty}^{\infty} \frac{dp_1}{2\pi} e^{-i\frac{p_1^2}{2m}\epsilon + ip_1(q_2 - q_1)}$

- $\langle q_1 | e^{-i\hat{H}\epsilon} | q_0 \rangle = e^{-iV(q_0)\epsilon} \int_{-\infty}^{\infty} \frac{dp_0}{2\pi} e^{-i\frac{p_0^2}{2m}\epsilon + ip_0(q_1 - q_0)}$

Note that we need N resolutions into momentum eigenstates instead of just $N - 1$: after inserting $N - 1$ resolutions of the identity into position eigenstates, we need one set of momentum eigenstates for each position "ket," $|q_0\rangle, ..., |q_{N-1}\rangle$, including the initial position q_0 over which we do not integrate. In any case, the amplitude is now:

$\langle q_F | e^{-i\hat{H}T} | q_0 \rangle =$

$\int_{-\infty}^{\infty} dq_1...dq_{N-1} \frac{dp_0}{2\pi}...\frac{dp_{N-1}}{2\pi} e^{-i\left[\frac{p_{N-1}^2}{2m}\epsilon - p_{N-1}(q_F - q_{N-1})\right]}...e^{-i\left[\frac{p_0^2}{2m}\epsilon - p_0(q_1 - q_0)\right]} e^{-iV(q_{N-1})}...e^{-iV(q_0)}$

Completing the square for each term in brackets will yield one-dimensional Gaussian integrals:

$$\frac{p_i^2}{2m}\epsilon - p_i(q_{i+1} - q_i) = \frac{\epsilon}{2m}\left[p_i^2 - \frac{2m}{\epsilon}p_i(q_{i+1} - q_i)\right]$$

$$= \frac{\epsilon}{2m}\left[\left(p_i - \frac{m}{\epsilon}(q_{i+1} - q_i)\right)^2 - \frac{m^2}{\epsilon^2}(q_{i+1} - q_i)^2\right]$$

$$= \frac{\epsilon}{2m}\left[\left(p_i - \frac{m}{\epsilon}(q_{i+1} - q_i)\right)^2\right] - \epsilon\frac{m}{2}\left(\frac{q_{i+1} - q_i}{\epsilon}\right)^2$$

Therefore, each of the momentum integrals gives:

$$\int_{-\infty}^{\infty} \frac{dp_i}{2\pi} e^{-i\left[\frac{p_i^2}{2m}\epsilon - p_i(q_{i+1} - q_i)\right]} = \frac{1}{2\pi}e^{+i\frac{m}{2}\left(\frac{q_{i+1} - q_i}{\epsilon}\right)^2} \int_{-\infty}^{\infty} dp_i\, e^{-i\frac{\epsilon}{2m}\left(p_i - \frac{m}{\epsilon}(q_{i+1} - q_i)\right)^2}$$

$$= \frac{1}{2\pi}e^{+i\frac{m}{2}\left(\frac{q_{i+1} - q_i}{\epsilon}\right)^2} \sqrt{\frac{\pi}{(i\epsilon/(2m))}}$$

$$= \sqrt{\frac{m}{2\pi i\epsilon}}\, e^{+i\frac{m}{2}\left(\frac{q_{i+1} - q_i}{\epsilon}\right)^2}$$

Again, we have N copies of this integral, so the transition amplitude is:

$$\langle q_F | e^{-i\hat{H}T} | q_0 \rangle = \int_{-\infty}^{\infty} dq_1...dq_{N-1} \left(\sqrt{\frac{m}{2\pi i\epsilon}}\right)^N e^{+i\frac{m}{2}\left(\frac{q_F - q_{N-1}}{\epsilon}\right)^2}...e^{+i\frac{m}{2}\left(\frac{q_1 - q_0}{\epsilon}\right)^2} e^{-iV(q_{N-1})}...e^{-iV(q_0)}$$

$$= \int d^{N-1}q \left(\sqrt{\frac{m}{2\pi i\epsilon}}\right)^N e^{i\sum_{j=0}^{N-1}\epsilon\left[\frac{m}{2}\left(\frac{q_{j+1} - q_j}{\epsilon}\right)^2 - V(q_j)\right]}$$

Remember that the limits $N \to \infty$ and $\epsilon \to 0$ with T fixed are supposed to be implied throughout. With this in mind, define the integral over paths and go to a continuum notation:

$$\int_{\substack{q(0) = q_0 \\ q(T) = q_F}} \mathcal{D}q \equiv \lim_{\substack{N \to \infty \\ \epsilon \to 0 \\ T \text{ fixed}}} \int d^{N-1}q \left(\sqrt{\frac{m}{2\pi i \epsilon}} \right)^N$$

$$\sum_{j=0}^{N-1} \epsilon \left[\frac{m}{2} \left(\frac{q_{j+1} - q_j}{\epsilon} \right)^2 - V(q_j) \right] \to \int_0^T dt \left[\frac{m}{2} \left(\frac{dq}{dt} \right)^2 - V(q) \right]$$

In the above, $q_N \equiv q_F$. This yields the desired result:

$$\langle q_F | e^{-iHT} | q_0 \rangle = \int \mathcal{D}q \, e^{i \int_0^T dt \left(\frac{1}{2} m \dot{q}^2 - V(q) \right)} \quad \text{with} \quad q(t = 0) = q_0 \, , \ q(t = T) = q_F$$

2. Derive (24)

$$\langle x_i x_j ... x_k x_\ell \rangle = \underset{\text{Wick}}{\sum} (A^{-1})_{ab} ... (A^{-1})_{cd} \qquad (24)$$

Solution:

Define the following N-dimensional integral:

$$\mathcal{Z}[\vec{J}] \equiv \int d^N x \, e^{-\frac{1}{2} \vec{x}^T M \vec{x} + \vec{J} \cdot \vec{x}} = \mathcal{C} e^{+\frac{1}{2} \vec{J}^T (M^{-1}) \vec{J}}$$

In the above, $\vec{x}$ and $\vec{J}$ are N-dimensional vectors, M is an $N \times N$ symmetric matrix, the superscript T denotes the transpose, and $\mathcal{C}$ is a constant that we set equal to 1. If you don't like that, then assume that all expectation values in what follows are implicitly divided by the constant $\mathcal{C}$. It is just a statistical normalization factor, the equivalent of ensuring that the sum of all probabilities equals 1 rather than some other number.

With this in mind, the expectation value of x_1 is:

$$\langle x_1 \rangle = \int d^N x \, e^{-\frac{1}{2} \vec{x}^T M \vec{x}} x_1$$

$$= \frac{\partial}{\partial J_1} \int d^N x \, e^{-\frac{1}{2} \vec{x}^T M \vec{x} + \vec{J} \cdot \vec{x}} \Big|_{\vec{J}=0}$$

$$= \frac{\partial}{\partial J_1} \mathcal{Z}[\vec{J}] \Big|_{\vec{J}=0}$$

4

In general, for an n-point function, we have:

$$\langle x_{i_1} x_{i_2} ... x_{i_n} \rangle = \frac{\partial}{\partial J_{i_1}} \frac{\partial}{\partial J_{i_2}} \cdots \frac{\partial}{\partial J_{i_n}} \mathcal{Z}[\vec{J}]\, |_{\vec{J}=0}$$

Take the case $n = 4$ to see how this works. To clean up the notation, define $G \equiv M^{-1}$ and use the following shorthand:

$$J \cdot G \cdot J \equiv \vec{J}^T G \vec{J} = \sum_{\alpha=1}^{N} \sum_{\beta=1}^{N} J_\alpha G_{\alpha\beta} J_\beta \equiv J_\alpha G_{\alpha\beta} J_\beta$$

First, we have the result from doing the Gaussian integral:

$$\mathcal{Z}[J] = e^{+\frac{1}{2} J \cdot G \cdot J}$$

Now take a derivative with respect to J_{i_4}:

$$\frac{\partial}{\partial J_{i_4}} e^{+\frac{1}{2} J \cdot G \cdot J} = e^{+\frac{1}{2} J \cdot G \cdot J} \frac{1}{2} \frac{\partial}{\partial J_{i_4}} \left(J_\alpha G_{\alpha\beta} J_\beta \right)$$

$$= e^{+\frac{1}{2} J \cdot G \cdot J} \frac{1}{2} \left(\delta_{i_4\alpha} G_{\alpha\beta} J_\beta + J_\alpha G_{\alpha\beta} \delta_{\beta i_4} \right)$$

$$= e^{+\frac{1}{2} J \cdot G \cdot J} G_{i_4\alpha_4} J_{\alpha_4}$$

The last line follows from the symmetry $G_{\alpha\beta} = G_{\beta\alpha}$. Also, we relabeled the dummy index to α_4 just to associate it with i_4 for later convenience. Before proceeding, note that setting $J = 0$ here gives zero, which means that $\langle x_{i_4} \rangle = 0$. Now take another derivative, this time with respect to J_{i_3}:

$$\frac{\partial^2}{\partial J_{i_3} \partial J_{i_4}} e^{+\frac{1}{2} J \cdot G \cdot J} = \frac{\partial}{\partial J_{i_3}} \left(e^{+\frac{1}{2} J \cdot G \cdot J} G_{i_4\alpha_4} J_{\alpha_4} \right)$$

$$= e^{+\frac{1}{2} J \cdot G \cdot J} \frac{\partial}{\partial J_{i_3}} \left(\frac{1}{2} J_\alpha G_{\alpha\beta} J_\beta \right) G_{i_4\alpha_4} J_{\alpha_4} + e^{+\frac{1}{2} J \cdot G \cdot J} G_{i_4\alpha_4} \frac{\partial}{\partial J_{i_3}} J_{\alpha_4}$$

$$= e^{+\frac{1}{2} J \cdot G \cdot J} \left(G_{i_3\alpha_3} J_{\alpha_3} G_{i_4\alpha_4} J_{\alpha_4} + G_{i_4 i_3} \right)$$

Again, before proceeding try setting $J = 0$. This time, a non-zero piece is left over. We see therefore that $\langle x_{i_3} x_{i_4} \rangle = G_{i_4 i_3} = (M^{-1})_{i_4 i_3}$. Now differentiate with respect to J_{i_2}:

$$\frac{\partial^3}{\partial J_{i_2} \partial J_{i_3} \partial J_{i_4}} e^{+\frac{1}{2} J \cdot G \cdot J} = \frac{\partial}{\partial J_{i_2}} \left[e^{+\frac{1}{2} J \cdot G \cdot J} \left(G_{i_3\alpha_3} J_{\alpha_3} G_{i_4\alpha_4} J_{\alpha_4} + G_{i_4 i_3} \right) \right]$$

$$= e^{+\frac{1}{2} J \cdot G \cdot J} \left(G_{i_2\alpha_2} J_{\alpha_2} \right) \left(G_{i_3\alpha_3} J_{\alpha_3} G_{i_4\alpha_4} J_{\alpha_4} + G_{i_4 i_3} \right) + e^{+\frac{1}{2} J \cdot G \cdot J} \left(G_{i_3 i_2} G_{i_4\alpha_4} J_{\alpha_4} + G_{i_3\alpha_3} J_{\alpha_3} G_{i_4 i_2} \right)$$

As in the case for just one derivative, setting $J = 0$ gives 0, which means that $\langle x_{i_2} x_{i_3} x_{i_4} \rangle = 0$. You can now see a general result for this generating function:

$$\langle x_{i_1} x_{i_2} \dots x_{i_n} \rangle = 0 \ \text{ if } n \text{ is odd}$$

Having taken enough derivatives to see how this works, while taking the next and final derivative, with respect to J_{i_1}, keep only the terms that have no powers of J left over, so that you don't have to bother keeping track of terms that will go to zero anyway. This gives:

$$\frac{\partial^4}{\partial J_{i_1} \partial J_{i_2} \partial J_{i_3} \partial J_{i_4}} e^{+\frac{1}{2} J \cdot G \cdot J} \big|_{J=0} = G_{i_2 i_1} G_{i_4 i_3} + G_{i_3 i_2} G_{i_4 i_1} + G_{i_3 i_1} G_{i_4 i_2}$$

On the right-hand side, all possible pairings of the indices $\{i_1, i_2, i_3, i_4\}$ appear exactly once.

I.3   From Mattress to Field

1. Verify that $D(x)$ decays exponentially for spacelike separation.

Solution:

The free-field propagator in $(3+1)$ spacetime dimensions is

$$D(x) = \int \frac{d^4 k}{(2\pi)^4} \frac{e^{ik \cdot x}}{k^2 - m^2 + i\varepsilon}$$

where $k \cdot x = k^0 x^0 - \vec{k} \cdot \vec{x}$. Rewriting this in the form

$$D(x^0, \vec{x}) = \int \frac{d^3 k}{(2\pi)^3} e^{-i\vec{k} \cdot \vec{x}} \int_{-\infty}^{\infty} \frac{dk^0}{2\pi} \frac{e^{ik^0 x^0}}{(k^0)^2 - (E_k - i\varepsilon)^2} \ , \quad E_k \equiv \sqrt{\vec{k}^2 + m^2}$$

we see that the integral over k^0 can be obtained by the contour integral

$$I \equiv \oint_C \frac{dz}{2\pi} \frac{e^{izx^0}}{z^2 - (E - i\varepsilon)^2}$$

with simple poles at $z = \pm(E - i\varepsilon)$. For $x^0 > 0$ we can use a semicircular contour in the upper half plane to pick up the residue at $z = -E + i\varepsilon$ and obtain

$$I = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \frac{e^{ik^0 x^0}}{(k^0)^2 - (E - i\varepsilon)^2} = i \lim_{z \to -E} \left(\frac{e^{izx^0}}{z - E} \right) = -i \frac{e^{-iEx^0}}{2E} \ , \quad x^0 > 0 \ .$$

For $x^0 < 0$ we can use a semicircular contour in the lower half plane to pick up the residue at $z = +E - i\varepsilon$ and, being careful to take into account the orientation of the contour, we obtain

$$I = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \frac{e^{ik^0 x^0}}{(k^0)^2 - (E - i\varepsilon)^2} = (-1)i \lim_{z \to +E} \left(\frac{e^{izx^0}}{z + E} \right) = -i \frac{e^{+iEx^0}}{2E}, \quad x^0 < 0 \, .$$

Therefore the propagator is

$$D(x^0, \vec{x}) = -i \int \frac{d^3k}{(2\pi)^3} \frac{1}{2E} e^{-iE|x^0| - i\vec{k} \cdot \vec{x}}, \quad E = \sqrt{\vec{k}^2 + m^2} \, .$$

Writing $d^3k = 2\pi \, dk \, k^2 \, d\cos\theta$ we can perform the integral over $\cos\theta$:

$$\int_{-1}^{1} d\cos\theta \, e^{-i\vec{k} \cdot \vec{x}} = \int_{-1}^{1} d\cos\theta \, e^{-ikr\cos\theta} = \frac{1}{-ikr} \left(e^{-ikr} - e^{+ikr} \right)$$

where $r \equiv |\vec{x}|$. We have

$$D(x^0, \vec{x}) = \frac{1}{8\pi^2 r} \int_0^{\infty} dk \, \frac{k}{\sqrt{k^2 + m^2}} e^{-i\sqrt{k^2 + m^2}|x^0|} \left(e^{-ikr} - e^{+ikr} \right)$$

Note that the integrand is even in k. Therefore

$$D(x^0, \vec{x}) = \frac{1}{16\pi^2 r} \int_{-\infty}^{\infty} dk \, \frac{k}{\sqrt{k^2 + m^2}} e^{-i\sqrt{k^2 + m^2}|x^0|} \left(e^{-ikr} - e^{+ikr} \right) \, .$$

Let $k \equiv m \sinh t$ so that $\sqrt{k^2 + m^2} = \cosh t$. We have

$$D(x^0, \vec{x}) = \frac{m}{16\pi^2 r} \int_{-\infty}^{\infty} dt \, \sinh t \, e^{-im|x^0|\cosh t} \left(e^{-imr\sinh t} - e^{+imr\sinh t} \right) \, .$$

From now on consider the spacelike coordinate vector $x^\mu = (0, \vec{x}) \implies x^\mu x_\mu = -r^2 < 0$. In other words, consider $D(r) \equiv D(0, \vec{x})$. The above integral with $x^0 = 0$ is almost in the form of the modified bessel functions, where the order-α modified bessel function of the second kind is defined as

$$K_\alpha(x) \equiv \frac{1}{2} e^{-i\frac{1}{2}\alpha\pi} \int_{-\infty}^{\infty} dt \, e^{-\alpha t - ix\sinh t} \, ,$$

which is a real function as long as x is real and positive. In our case, we have $x = mr$ and therefore

$$D(r) = -i \frac{m}{4\pi^2 r} K_1(mr) \, .$$

For $mr \gg 1$, the function $K_1(mr)$ behaves as

$$K_1(mr) \approx \sqrt{\frac{\pi}{2mr}} e^{-mr}$$

showing that $D(r \to \infty) \sim e^{-mr} \to 0$. Alternatively, for $m \to 0$ we have

$$K_1(mr) \approx \frac{1}{mr} \implies D(r) \approx -i \frac{1}{4\pi^2 r^2} \, .$$

7

2. Work out the propagator $D(x)$ for a free field theory in $(1+1)$-dimensional spacetime and study the large x^1 behavior for $x^0 = 0$.

Solution:

The propagator in $(1+1)$ spacetime dimensions is

$$D(x) = \int \frac{d^2 k}{(2\pi)^2} \frac{e^{ik \cdot x}}{k^2 - m^2 + i\varepsilon} .$$

Performing the integral over k^0 and setting $x^0 = 0$ as in problem I.3.1 gives

$$D(0, x) = -i\frac{1}{4\pi} \int_{-\infty}^{\infty} dk \frac{1}{\sqrt{k^2 + m^2}} e^{-ikx} .$$

The substitution $k \equiv m \sinh t$ gives

$$D(0, x) = -i\frac{1}{4\pi} \int_{-\infty}^{\infty} dt\, e^{-imx \sinh t} = -i\frac{1}{2\pi} K_0(mr)$$

where we have used the definition of the modified bessel function from problem I.3.1. For $mr \gg 1$, we obtain

$$D(x, 0) \approx -i\frac{1}{2\pi} \left(\sqrt{\frac{\pi}{2mr}} e^{-mr} \right) \sim e^{-mr}$$

as in $(3+1)$ dimensions. In contrast to $(3+1)$ dimensions, for $m \to 0$ we have

$$K_0(mr) \approx -\ln(mr/2) - \gamma$$

where $\gamma \approx 0.577$.

3. Show that the advanced propagator defined by

$$D_{\text{adv}}(x - y) = \int \frac{d^4 k}{(2\pi)^4} \frac{e^{ik(x-y)}}{k^2 - m^2 - i \operatorname{sgn}(k_0)\varepsilon}$$

is nonzero only if $x^0 > y^0$. In other words, it only propagates into the future. [Hint: Both poles of the integrand are now in the upper half of the k_0-plane.] Incidentally, some authors prefer to write $(k_0 - i\varepsilon)^2 - \vec{k}^2 - m^2$ instead of $k^2 - m^2 - i\operatorname{sgn}(k_0)\varepsilon$ in the integrand. Similarly, show that the retarded propagator

$$D_{\text{ret}}(x - y) = \int \frac{d^4 k}{(2\pi)^4} \frac{e^{ik(x-y)}}{k^2 - m^2 + i \operatorname{sgn}(k_0)\varepsilon}$$

propagates into the past.

8

Solution:

The advanced propagator is:

$$D_{\text{adv}}(x) = \int \frac{d^4k}{(2\pi)^4} \frac{e^{ik_\mu x^\mu}}{k_\nu k^\nu - m^2 - i\operatorname{sgn}(k^0)\varepsilon}$$

$$= \int \frac{d^3k}{(2\pi)^3} \int_{-\infty}^{\infty} \frac{dk^0}{2\pi} \frac{e^{ik^0 x^0} e^{-i\vec{k}\cdot\vec{x}}}{(k^0)^2 - |\vec{k}|^2 - m^2 - i\operatorname{sgn}(k^0)\varepsilon}$$

$$= \int \frac{d^3k}{(2\pi)^3} e^{-i\vec{k}\cdot\vec{x}} \int_{-\infty}^{\infty} \frac{dk^0}{2\pi} \frac{e^{ik^0 x^0}}{\left[k^0 - (\omega_{\vec{k}} + i\operatorname{sgn}(k^0)\varepsilon)\right]\left[k^0 + (\omega_{\vec{k}} + i\operatorname{sgn}(k^0)\varepsilon)\right]}$$

In the above, $\omega_{\vec{k}} \equiv \sqrt{|\vec{k}|^2 + m^2}$, $O(\varepsilon^2)$ terms are set to zero, and as explained in the book we take ε to be a generic positive infinitesimal, so for instance we write "$2\varepsilon = \varepsilon$" instead of bothering to define new symbols.

Now evaluate this integral using the complex plane. Define the complex variable $z \equiv k^0 + i\beta$, where β is real, and consider the following integral:

$$\mathcal{I}(x^0) \equiv \int_c \frac{dz}{2\pi} \frac{e^{izx^0}}{\left[z - (\omega + i\operatorname{sgn}(\operatorname{Re}\{z\}))\right]\left[z + (\omega + i\operatorname{sgn}(\operatorname{Re}\{z\}))\right]}$$

So far the contour c is unspecified. The intention is to provide a contour such that integration along the real axis will yield the $\int dk^0$ that we're actually interested in evaluating, and that the contributions along the imaginary parts of the complex plane will evaluate to zero.

In anticipation of an appropriate choice of contour, consider the behavior of the integrand in $\mathcal{I}(x^0)$ when z is pure-imaginary: $z \to i\beta$, where β is real. The exponential factor in the numerator becomes $e^{-\beta x^0}$. If $x^0 > 0$, then this exponential goes to zero for $\beta \to +\infty$. If $x^0 < 0$, then this exponential goes to zero for $\beta \to -\infty$. The denominator of the integrand is a polynomial, so if the exponential goes to zero then the whole integrand will go to zero.

So now we've distinguished the two cases: for $x^0 > 0$, we must close the contour in the upper half-plane, and for $x^0 < 0$, we must close the contour in the lower half-plane. The relevant contours for this problem, which will yield the desired real-valued integral, are shown below:

9
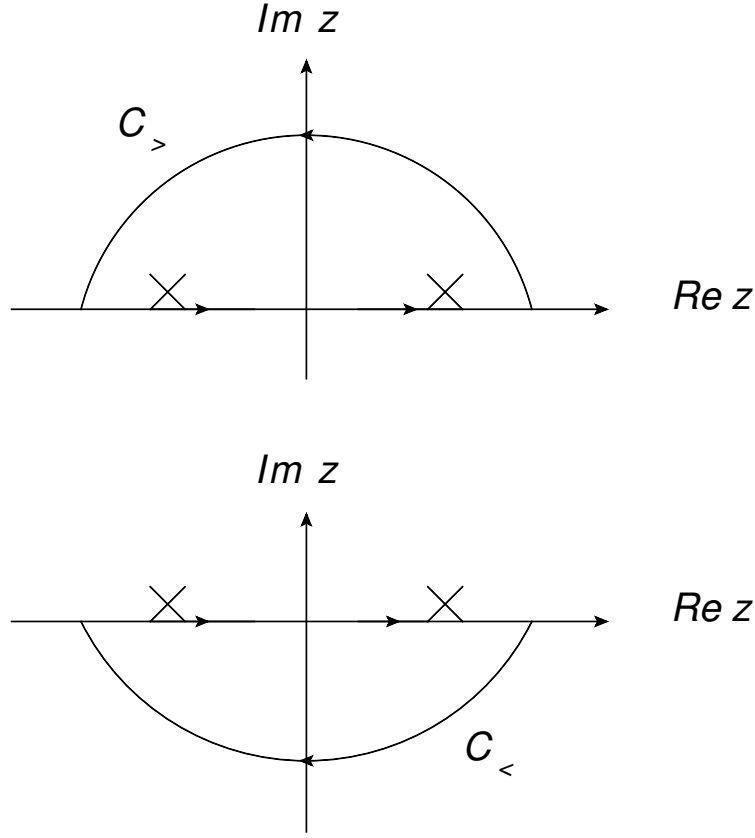
$c_>$ is what we use for $x^0 > 0$, and $c_<$ is what we use for $x^0 < 0$. If you evaluate the integrand for the piece of either $c_>$ or $c_<$ that is on the real axis, you will recover the integral you are actually interested in evaluating ($\text{Re}\{z\} = k^0$). The $\times$s in the above diagram indicate the locations of the poles of the integrand. A pole is where the integrand goes to infinity, which in this case corresponds to a zero of the denominator. The pole for $\text{Re}\{z\} > 0$ is $z = \omega + i\,\text{sgn}(\text{Re}\{z\})\varepsilon = \omega + i\varepsilon$. The pole for $\text{Re}\{z\} < 0$ is $z = -\omega - i\,\text{sgn}(\text{Re}\{z\})\varepsilon = -\omega + i\varepsilon$. So both of the poles have positive imaginary parts and are therefore in the upper half-plane.

The residue theorem says that if you choose a contour that encircles the locations of the poles of the integrand, then you will get some nonzero number for the integral. If, however, your contour does not encircle any of the poles, then you will get zero. For the purposes of this problem all you have to do is decide whether the integral is zero.

From the above discussion, we have determined two facts:

1. All poles of the integrand are in the upper half-plane.

2. If $x^0 > 0$, we close the contour in the upper half-plane. If $x^0 < 0$, we close in the lower half-plane.

Therefore, the integral is nonzero if $x^0 > 0$, and the integral is zero if $x^0 < 0$. Therefore, the advanced propagator is only nonzero if $x^0 > 0$.

10

For the retarded propagator, everything is the same except the poles of the integrand are in the lower half-plane. So the retarded propagator is only nonzero if we close the contour in the lower half-plane, which we can only do when $x^0 < 0$.

I.4 From Field to Particle to Force

1. Calculate the analog of the inverse square law in a (2+1)-dimensional universe, and more generally in a (D+1)-dimensional universe.

Solution:

Taking the analog of equation (1) on page 24 of the book, we have in $(d+1)$ dimensions:

$$Z[J] \propto e^{iW[J]} \; , \; W[J] = -\frac{1}{2} \int \int d^{d+1}x \, d^{d+1}y \, J(x) D(x-y) J(y)$$

The free-field propagator $D(x-y)$ is:

$$D(x-y) = \int \frac{d^{d+1}k}{(2\pi)^{d+1}} \frac{-e^{ik(x-y)}}{-k^2 + m^2 - i\epsilon}$$

As in the book, consider $J(x^0, \vec{x}) = \delta^{(d)}(\vec{x} - \vec{x}_1) + \delta^{(d)}(\vec{x} - \vec{x}_2)$, and look only at the cross term $\sim \delta^{(d)}(\vec{x} - \vec{x}_1) \delta^{(d)}(\vec{x} - \vec{x}_2)$ in $W[J]$:

$W[J]_{\text{cross}}$

$$= -\int dx^0 \int d^d x \int dy^0 \int d^d y \int \frac{d^d k}{(2\pi)^d} \int \frac{dk^0}{2\pi} \frac{-e^{ik^0(x^0-y^0)} e^{-i\vec{k} \cdot (\vec{x} - \vec{y})}}{-(k^0)^2 + |\vec{k}|^2 + m^2 - i\epsilon} \delta^{(d)}(\vec{x} - \vec{x}_1) \delta^{(d)}(\vec{x} - \vec{x}_2)$$

$$= -\int dx^0 \int dy^0 \int \frac{d^d k}{(2\pi)^d} \int \frac{dk^0}{2\pi} \frac{-e^{ik^0(x^0-y^0)} e^{-i\vec{k} \cdot (\vec{x}_1 - \vec{x}_2)}}{-(k^0)^2 + |\vec{k}|^2 + m^2 - i\epsilon}$$

$$= -\int dx^0 \int \frac{d^d k}{(2\pi)^d} \int \frac{dk^0}{2\pi} \underbrace{\left(\int dy^0 \, e^{-ik^0 y^0} \right)}_{2\pi \delta(k^0)} \frac{-e^{ik^0 x^0} e^{-i\vec{k} \cdot (\vec{x}_1 - \vec{x}_2)}}{-(k^0)^2 + |\vec{k}|^2 + m^2 - i\epsilon}$$

$$= -\left(\int dx^0 \right) \int \frac{d^d k}{(2\pi)^d} \frac{-e^{-i\vec{k} \cdot (\vec{x}_1 - \vec{x}_2)}}{|\vec{k}|^2 + m^2 - i\epsilon}$$

$$= +T \int \frac{d^d k}{(2\pi)^d} \frac{e^{-i\vec{k} \cdot (\vec{x}_1 - \vec{x}_2)}}{|\vec{k}|^2 + m^2}$$

Since $iW = -iET$, we have an expression for the potential energy:

$$E = -\int \frac{d^d k}{(2\pi)^d} \frac{e^{-i\vec{k} \cdot (\vec{x}_1 - \vec{x}_2)}}{|\vec{k}|^2 + m^2}$$

11

Now set m to zero and use dimensional analysis:

$$E \sim \int dk \, k^{d-1} \frac{e^{ikr}}{k^2} = \int dk \, k^{d-3} \, e^{ikr} = \int \frac{du}{r} \left(\frac{u}{r}\right)^{d-3} e^{iu} \sim \frac{1}{r^{d-2}}$$

The notation is intended to extract the dimensional information only; for instance, $\vec{k} \cdot \vec{x} = kr \cos\theta$, etc, but for the purposes of this problem the angular information is not important. For $d = 3$, the above gives $E \sim 1/r$, as expected.

However, for $d = 2$ this gives $E \sim r^0 = 1$, which may naively suggest that the force between the sources is zero, but that is incorrect. Define $r \equiv |\vec{x}_1 - \vec{x}_2|$ and consider the force between the sources:

$$
\begin{aligned}
F &= -\frac{dE}{dr} = +\frac{d}{dr} \int \frac{d^2k}{(2\pi)^2} \frac{e^{-i\vec{k}\cdot(\vec{x}_1 - \vec{x}_2)}}{|\vec{k}|^2} \\
&= \frac{1}{4\pi^2} \frac{d}{dr} \int_0^\infty k \, dk \int_0^{2\pi} d\theta \frac{e^{-ikr\cos\theta}}{k^2} \\
&= \frac{1}{4\pi^2} \int_0^\infty \frac{dk}{k} \int_0^{2\pi} d\theta \, e^{-ikr\cos\theta} \, (-ik\cos\theta) \\
&= \frac{-i}{4\pi^2 r} \int_0^\infty du \int_0^{2\pi} d\theta \, e^{-iu\cos\theta} \, \cos\theta \\
&\sim \frac{1}{r}
\end{aligned}
$$

The force is not zero; it goes as $1/r$. The potential energy is therefore $E \sim \ln r$. However this is still not quite right, since the argument of the log must be dimensionless. So really we get $E \sim \ln(r/r_0)$ for some cutoff r_0, which explains our naive estimate that the potential energy is dimensionless. Let's study the properties of this cutoff.

Explicitly, the potential energy is:

$$E(r) = -\frac{1}{4\pi^2} \int_0^\infty \frac{dk}{k} \int_0^{2\pi} d\theta \, e^{-ikr\cos\theta}$$

The integral $\frac{1}{2\pi} \int_0^{2\pi} d\theta \, e^{-ikr\cos\theta} = J_0(kr)$, where $J_0(z)$ is the Bessel function of the first kind of order 0 as a function of z. We may now use Mathematica's Bessel function input $\mathrm{BesselJ}[n, z] \equiv J_n(z)$ to plot the integral. Plotting the dimensionless function $\int du \, J_0(u)/u$ gives:

12

$\int du \, J_0(u)/u$
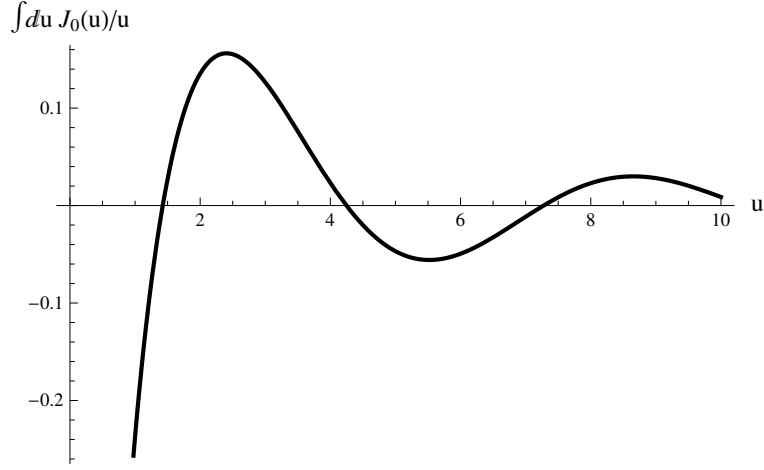
Identifying $u = kr$, we see that the integral over k is uneventful at high k but diverges as $k \to 0$. This suggests we regularize the integral by imposing a cutoff $k_{\min} \equiv 1/a$ on the lower bound of integration. So really, the potential energy is:

$$E(r) = -\frac{1}{4\pi} \int_{1/a}^{\infty} \frac{dk}{k} \int_0^{2\pi} d\theta \, e^{-ikr\cos\theta}$$

Remember that in our units of $\hbar = c = 1$, momentum has units of inverse length. If we take the potential energy as written above and change variables to $u = kr$, then the potential energy is:

$$E(r) = -\frac{1}{4\pi} \int_{r/a}^{\infty} \frac{du}{u} \int_0^{2\pi} d\theta \, e^{-iu\cos\theta}$$

Now we see that the expression for the potential energy is more sensible: the reason we naively guessed that E does not depend on r is that we had improperly chosen the lower bound of the k integral to be 0. We also now see what the r_0 is in $E \sim \ln(r/r_0)$: $r_0 = a = 1/k_{\min}$ is the inverse of the minimum momentum, or equivalently the maximum separation between the particles that we are willing to consider.

Finally, the double integral in the force is $\int_0^{\infty} du \int_0^{2\pi} d\theta \, e^{-iu\cos\theta} \cos\theta = -2\pi i$, so the force between the particles is

$$F = -\frac{1}{2\pi r} \ .$$

13

I.5 Coulomb and Newton: Repulsion and Attraction

1. Write down the most general form for $\sum_a \varepsilon_{\mu\nu}^{(a)}(k)\varepsilon_{\lambda\sigma}^{(a)}(k)$ using symmetry repeatedly. For example, it must be invariant under the exchange $\{\mu\nu \leftrightarrow \lambda\sigma\}$. You might end up with something like

$$
AG_{\mu\nu}G_{\lambda\sigma} + B(G_{\mu\lambda}G_{\nu\sigma} + G_{\mu\sigma}G_{\nu\lambda}) + C(G_{\mu\nu}k_\lambda k_\sigma + k_\mu k_\nu G_{\lambda\sigma})
$$
$$
+ D(k_\nu k_\lambda G_{\nu\sigma} + k_\mu k_\sigma G_{\nu\lambda} + k_\nu k_\sigma G_{\mu\lambda} + k_\nu k_\lambda G_{\mu\sigma}) + Ek_\mu k_\nu k_\lambda k_\sigma
$$

with various unknown $A, ..., E$. Apply $k^\mu \sum_a \varepsilon_{\mu\nu}^{(a)}(k)\varepsilon_{\lambda\sigma}^{(a)}(k) = 0$ and find out what that implies for the constants. Proceeding in this way, derive (13).

$$
(13) \qquad \sum_a \varepsilon_{\mu\nu}^{(a)}(k)\varepsilon_{\lambda\sigma}^{(a)}(k) = \frac{(G_{\mu\lambda}G_{\nu\sigma} + G_{\mu\sigma}G_{\nu\lambda}) - \frac{2}{3}G_{\mu\nu}G_{\lambda\sigma}}{k^2 - m^2}
$$

Solution:

The ingredients we have are k_μ and $G_{\mu\nu}(k) \equiv g_{\mu\nu} - \frac{1}{m^2}k_\mu k_\nu$. G has the properties that $G_{\mu\nu} = G_{\nu\mu}$ and $k^\mu G_{\mu\nu} = k^\mu\left(g_{\mu\nu} - \frac{1}{m^2}k_\mu k_\nu\right) = k_\nu - \frac{1}{m^2}(m^2)k_\nu = 0$.

Each polarization tensor $\varepsilon_{\mu\nu}^{(a)}$ is symmetric under interchange of μ and ν, so we need $f_{\mu\nu\lambda\sigma} = f_{\nu\mu\lambda\sigma}$ and $f_{\mu\nu\lambda\sigma} = f_{\mu\nu\sigma\lambda}$. This leads us to write:

$$
f_{\mu\nu\lambda\sigma} \supset A\,G_{\mu\nu}G_{\lambda\sigma} + C\,G_{\mu\nu}k_\lambda k_\sigma + \tilde{C}\,k_\mu k_\nu G_{\lambda\sigma}
$$

However, because the components of the polarization tensors are just ordinary numbers, we know that $\varepsilon_{\mu\nu}\varepsilon_{\lambda\sigma} = \varepsilon_{\lambda\sigma}\varepsilon_{\mu\nu}$, which means that f should be invariant under the interchange of pairs of indices, $(\mu\nu) \leftrightarrow (\lambda\sigma)$: $f_{\mu\nu\lambda\sigma} = f_{\lambda\sigma\mu\nu}$. This sets $\tilde{C} = C$.

Another thing we can do is stagger the indices across Gs, by which I mean something involving the term $G_{\mu\lambda}G_{\nu\sigma}$. If we take that and interchange μ and ν, we get $G_{\nu\lambda}G_{\mu\sigma}$. We therefore have another term to add to f:

$$
f_{\mu\nu\lambda\sigma} \supset B\,(G_{\mu\lambda}G_{\nu\sigma} + G_{\nu\lambda}G_{\mu\sigma})
$$

We can also use the same idea but instead of using two Gs, use two ks and one G, such as $k_\mu k_\lambda G_{\nu\sigma}$. If we take that and interchange μ and ν, we get $k_\nu k_\lambda G_{\mu\sigma}$. If we then interchange λ and σ, we get $k_\nu k_\sigma G_{\mu\lambda}$. Interchanging μ and ν again gives us something different: $k_\mu k_\sigma G_{\nu\lambda}$. These four terms should be included in f:

$$
f_{\mu\nu\lambda\sigma} \supset D\,(k_\mu k_\lambda G_{\nu\sigma} + k_\nu k_\lambda G_{\mu\sigma} + k_\nu k_\sigma G_{\mu\lambda} + k_\mu k_\sigma G_{\nu\lambda})
$$

Finally, there is the possibility of multiplying together four momenta to yield a term $f_{\mu\nu\lambda\sigma} \supset$

$E\, k_\mu k_\nu k_\lambda k_\sigma$. We therefore have:

$$f_{\mu\nu\lambda\sigma} = A\, G_{\mu\nu} G_{\lambda\sigma} + B\, (G_{\mu\lambda} G_{\nu\sigma} + G_{\nu\lambda} G_{\mu\sigma}) + C\, (G_{\mu\nu} k_\lambda k_\sigma + k_\mu k_\nu G_{\lambda\sigma})$$
$$+ D\, (k_\mu k_\lambda G_{\nu\sigma} + k_\nu k_\lambda G_{\mu\sigma} + k_\nu k_\sigma G_{\mu\lambda} + k_\mu k_\sigma G_{\nu\lambda}) + E\, k_\mu k_\nu k_\lambda k_\sigma$$

Now we need to fix the constants A, B, C, D and E. To do that, we use the fact that $k^\mu \varepsilon_{\mu\nu}(k) = 0$, which implies that $k^\mu f_{\mu\nu\lambda\sigma} = 0$. Recalling that $k^\mu G_{\mu\nu} = 0$ and $k^\mu k_\mu = m^2$ gives the following:

$$k^\mu f_{\mu\nu\lambda\sigma} = A(0) + B(0+0) + C(0 + m^2 k_\nu G_{\lambda\sigma}) + D(m^2 k_\lambda G_{\nu\sigma} + 0 + 0 + m^2 k_\sigma G_{\nu\lambda}) + Em^2 k_\nu k_\lambda k_\sigma$$
$$= 0 \implies C k_\nu G_{\lambda\sigma} + D(k_\lambda G_{\nu\sigma} + k_\sigma G_{\nu\lambda}) = 0 \ , \quad E = 0$$

Multiplying by k^ν implies $C = 0$, and multiplying by k^λ implies $D = 0$. We are now left with:

$$f_{\mu\nu\lambda\sigma} = A\, G_{\mu\nu} G_{\lambda\sigma} + B\, (G_{\mu\lambda} G_{\nu\sigma} + G_{\nu\lambda} G_{\mu\sigma})$$

To proceed we recall the other trace-free condition on the polarization tensor, namely that $g^{\mu\nu} \varepsilon_{\mu\nu} = 0$. In anticipation of multiplying $f_{\mu\nu\lambda\sigma}$ by $g^{\mu\nu}$, we first compute the following:

$$g^{\mu\nu} G_{\mu\lambda} = g^{\mu\nu} \left(g_{\mu\lambda} - \frac{1}{m^2} k_\mu k_\lambda \right)$$
$$= \delta^\nu_\lambda - \frac{1}{m^2} k^\nu k_\lambda$$
$$\implies g^{\mu\nu} G_{\mu\nu} = \delta^\nu_\nu - \frac{1}{m^2} k^\nu k_\nu = 4 - 1 = 3$$

Therefore, the second trace-free condition implies:

$$g^{\mu\nu} f_{\mu\nu\lambda\sigma} = A(3) G_{\lambda\sigma} + B \left[\left(\delta^\nu_\lambda - \frac{1}{m^2} k^\nu k_\lambda \right) G_{\nu\sigma} + G_{\nu\lambda} \left(\delta^\nu_\sigma - \frac{1}{m^2} k^\nu k_\sigma \right) \right]$$
$$= 3A\, G_{\lambda\sigma} + B \left[\left(\delta^\nu_\lambda - \frac{1}{m^2} k^\nu k_\lambda \right) \left(g_{\nu\sigma} - \frac{1}{m^2} k_\nu k_\sigma \right) + \left(g_{\nu\lambda} - \frac{1}{m^2} k_\nu k_\lambda \right) \left(\delta^\nu_\sigma - \frac{1}{m^2} k^\nu k_\sigma \right) \right]$$
$$= 3A\, G_{\lambda\sigma} + 2B \left[g_{\lambda\sigma} - \frac{1}{m^2} k_\lambda k_\sigma - \frac{1}{m^2} k_\sigma k_\lambda + \left(\frac{1}{m^2} \right)^2 (m^2) k_\lambda k_\sigma \right]$$
$$= 3A\, G_{\lambda\sigma} + 2B\, G_{\lambda\sigma} = 0 \implies A = \frac{-2B}{3}$$

We now have the answer up to an overall normalization constant:

15

$$f_{\mu\nu\lambda\sigma} = B\left[-\frac{2}{3}G_{\mu\nu}G_{\lambda\sigma} + (G_{\mu\lambda}G_{\nu\sigma} + G_{\lambda\mu}G_{\sigma\nu})\right]$$

As suggested in the book, fix the constant B by imposing the normalization $f_{1212} = 1$ for all k. If this is supposed to hold for all k, then try $k = 0$. Since $G_{ij}(k = 0) = 1$ for $i, j = 1, 2, 3$, we get:

$$1 = f_{1212}(0) = B\, G_{11}(0)G_{22}(0) = B \implies B = 1$$

We now have the desired result:

$$f_{\mu\nu\lambda\sigma} = -\frac{2}{3}G_{\mu\nu}G_{\lambda\sigma} + (G_{\mu\lambda}G_{\nu\sigma} + G_{\lambda\mu}G_{\sigma\nu})$$

I.6 Inverse Square Law and the Floating 3-Brane

1. Putting in the numbers, show that the case $n = 1$ is already ruled out.

$$M^2_{\text{Pl}(n+3+1)} = \frac{M^2_{\text{Pl}}}{[M_{\text{Pl}(n+3+1)}R]^n}$$

Solution:

To clarify what this question is asking: The original point of this compact extra dimension setup is to propose that the scale of quantum gravity in the $(n+3+1)$-dimensional space is actually similar to the weak scale, so that our guessing that the Planck mass is so high is just an artifact of measuring in $(3+1)$ dimensions.

So, really, we are interested in answering the question: "Can one extra dimension fix the hierarchy problem?"

With that in mind, we set $n = 1$ and try $M_{\text{Pl}(5)} \sim 10^2$ GeV. Using $M_{\text{Pl}} \sim 10^{19}$ GeV, we predict a value for the characteristic length scale of the proposed 5th extra dimension:

$$\frac{1}{R} = M_{\text{Pl}(5)}\left(\frac{M_{\text{Pl}(5)}}{M_{\text{Pl}}}\right)^2 \sim 10^2\,\text{GeV}\left(\frac{10^2}{10^{19}}\right)^2 = 10^{-32}\,\text{GeV} = 10^{-23}\,\text{eV} \sim 10^{-42}\,\text{J}$$

Remembering that $\hbar c \sim 10^{-26}$ J·m gets us from $\hbar = c = 1$ units to SI units, we get:

$$R \to \frac{R}{\hbar c} \sim \frac{10^{42}}{\text{J}} \implies R \sim \frac{10^{42}}{\text{J}}\left(10^{-26}\,\text{J}\cdot\text{m}\right) = 10^{16}\,\text{m} \sim 1\,\text{parsec}$$

We conclude that adding only one extra dimension is not a valid way to fix the hierarchy problem.

I.7   Feynman Diagrams

1. Work out the amplitude corresponding to figure I.7.11 in (24).

Solution:

Since this is a 3-loop amplitude, we will just write down the answer using the Feynman rules and then get the symmetry factor by looking at the diagram. Using the propagator $i\Delta(p) = \frac{i}{p^2 - m^2 + i\varepsilon}$ and the vertex $-i\lambda$, we have

$$\mathcal{M} =$$
$$\frac{1}{S}\int \frac{d^4p}{(2\pi)^4}\frac{d^4q}{(2\pi)^4}\frac{d^4r}{(2\pi)^4}[-i\lambda]^4[i\Delta(p)][i\Delta(k_1+k_2-p)][i\Delta(q)][i\Delta(p-q-r)][i\Delta(k_1+k_2-r)][i\Delta(r)]$$

where S is the symmetry factor. Up to the symmetry factor, this matches (24). Symmetry factors arise in loop diagrams for the following reason. Consider the pictorial representation of the vertices and of the propagators with sources:
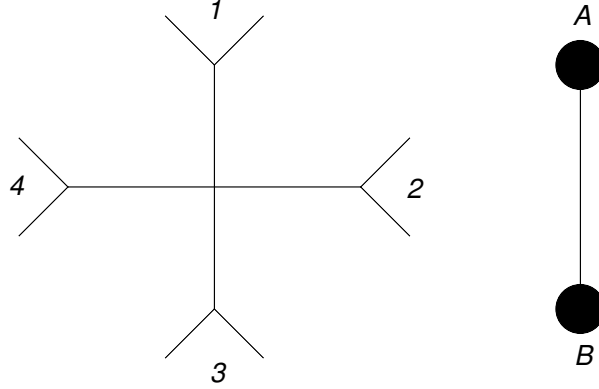
We normalize the vertex as $\mathcal{L} = -\frac{1}{4!}\lambda\varphi^4$ so that the permutations of the labels $1, 2, 3, 4$ on the "hands" contribute 4! terms and cancel out the $\frac{1}{4!}$ in the Lagrangian. We normalize the kinetic term as $\mathcal{L} = -\frac{1}{2}\varphi(\partial^2 + m^2)\varphi$ so that swapping the labels A, B on the "blobs" contribute 2 terms and cancel out the $\frac{1}{2}$ in the Lagrangian.

Suppose we have a diagram in which hands 1 and 2 eat blobs A and B respectively:
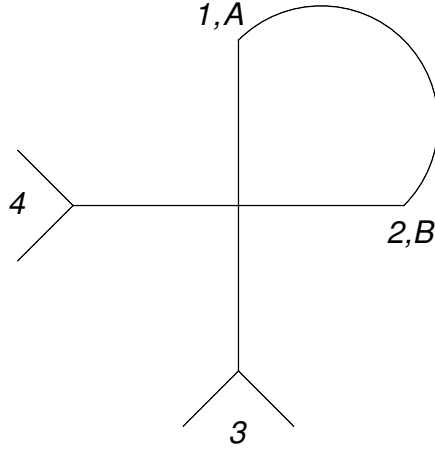
From our previous argument, this diagram takes into account the 2 terms from swapping hands 1 and 2, and also the 2 terms from swapping blobs A and B, leading to a total of 4 terms.

But swapping 1 and 2 and simultaneously swapping A and B gives exactly the same diagram with which we started, so this bookkeeping overcounts by a factor of $S = 2$.

Now instead of rearranging individual hands on a given vertex, and individual blobs on a given propagator, consider rearranging the vertices themselves and the propagators themselves. Consider the path integral

$$\mathcal{Z}(J) = \int \mathcal{D}\varphi \, e^{i\int d^4x(\mathcal{L}+J\varphi)}$$

$$= \sum_{V=0}^{\infty} \frac{1}{V!} \left[-i \int d^4x \frac{1}{4!} \lambda \left(\frac{1}{i} \frac{\delta}{\delta J(x)} \right)^4 \right]^V \sum_{P=0}^{\infty} \frac{1}{P!} \left[i \int d^4x \, d^4y \frac{1}{2} J(x)\Delta(x-y)J(y) \right]^P .$$

The quantity in the first square brackets raised to the power V is the vertex diagram drawn above, and the quantity in the second square brackets raised to the power P is the barbell diagram drawn above.

Thus we see (as explained in the main text) that a general diagram is given by pasting together a bunch of different vertices with a bunch of different propagator lines. The blobs (the $J(x)$) on the propagator lines get eaten by the hands (the $\frac{\delta}{\delta J(x)}$) on the vertices.

Just as before, we might think that the V vertices can be arranged in $V!$ ways, and the P propagators can be arranged in $P!$ ways, so that the $V!P!$ cancel out of the path integral. But if exchanging two vertices and simultaneously exchanging two propagators gives you the same diagram you started with, then counting those two swaps separately would overcount the contribution of that term to the sum in $\mathcal{Z}(J)$.

The mismatch between the naive counting the correct counting of matching derivatives to sources is the symmetry factor. The last thing to take note of is that symmetry factors

coming from external blobs are canceled once those blobs are given fixed external values. In other words, tree diagrams do not have symmetry factors.

Now let us apply this reasoning to figure I.7.11, which we repeat below for convenience:
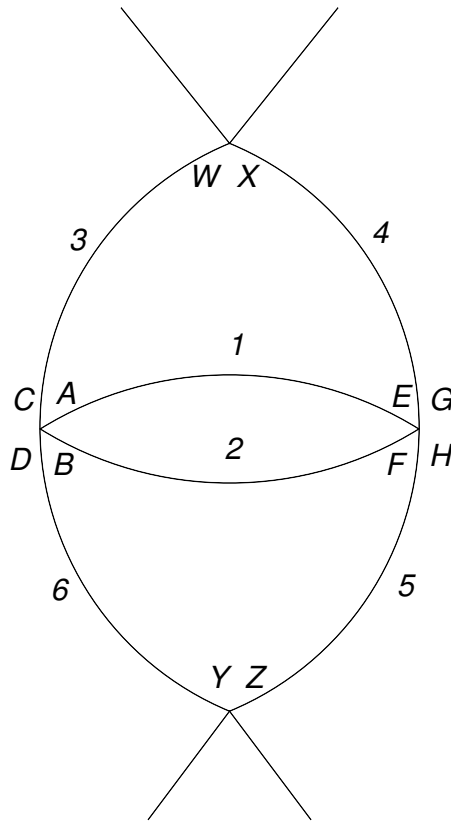
Here we omit the momentum arrows and labels, since they play no role in counting the symmetry factor.[1] We label internal lines by numbers, and "hands" on the vertices by letters.

Suppose we swap the vertices labeled by $(ABCD)$ and $(EFGH)$. Can we swap the internal lines in such a way that we recover the original diagram? The answer is yes: swap lines $3 \leftrightarrow 4$ and $5 \leftrightarrow 6$, and then also the hands $W \leftrightarrow X$ and $Y \leftrightarrow Z$. This contributes a factor of 2 to S.

This is also the same as keeping the vertices fixed, but swapping the hands $A \leftrightarrow B$ and $E \leftrightarrow F$, while also swapping the lines $1 \leftrightarrow 2$. This contributes another factor of 2 to S.

There is nothing else we can do, so we find $S = 2 \times 2 = 4$.

_____

[1]In theories with complex scalar fields, the arrows can denote the flow of a conserved $U(1)$ charge. In such cases, the arrows do play a role in determining the symmetry factor since the $U(1)$ charge must remain conserved.

3. Draw all the diagrams describing two mesons producing four mesons up to and including order λ^2. Write down the corresponding Feynman amplitudes.

Solution:

The Lagrangian density is:

$$\mathcal{L} = \frac{1}{2}\left((\partial\varphi)^2 - m^2\,\varphi^2\right) - \frac{1}{4!}\,\lambda\,\varphi^4$$

This yields the following propagator and vertex:
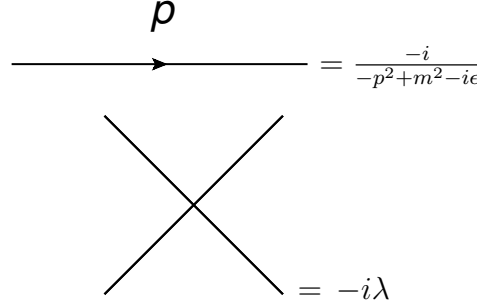
The $2 \to 4$ scattering diagrams up to $O(\lambda^2)$ are:
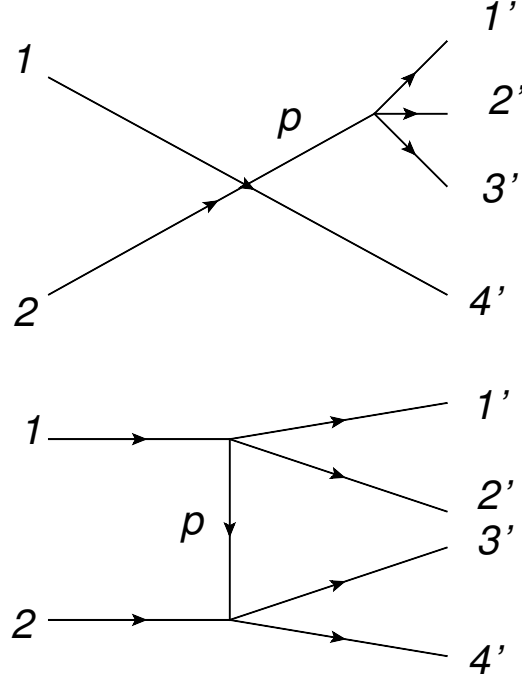
There are also the diagrams with different permutations of the outgoing lines, which correspond to different momenta flowing in the internal propagator. Reading the first diagram from right to left, the corresponding amplitude is:

$$i\mathcal{M}_{\text{first}} = [-i\lambda]\left[\frac{-i}{-p^2 + m^2 - i\epsilon}\right][-i\lambda] = +i\,\frac{\lambda^2}{-p^2 + m^2 - i\epsilon}$$

20

In the first diagram, the propagator's momentum is $p = p_1 + p_2 - p_{4'}$. The amplitude for the second diagram has the same form as the first one, except $p = p_1 - (p_{1'} + p_{2'})$. So the total amplitude is, suppressing the $i\varepsilon$s:

$$
\mathcal{M} = \lambda^2 \left[\sum_{n=1'}^{4'} \frac{1}{-(p_1 + p_2 - p_{n'})^2 + m^2} + \sum_{\substack{i=1' \\ i'<j'}}^{4'} \sum_{j=1'}^{4'} \frac{1}{-(p_1 - p_{i'} - p_{j'})^2 + m^2} \right] + O(\lambda^4) \ .
$$

There is no symmetry factor since this is a tree diagram.

I.8 Quantizing Canonically

1. Derive (14). Then verify explicitly that $d^D k/(2\omega_k)$ is indeed Lorentz invariant. Some authors prefer to replace $\sqrt{2\omega_k}$ in (11) by $2\omega_k$ when relating the scalar field to the creation and annihilation operators. Show that the operators defined by these authors are Lorentz covariant. Work out their commutation relation.

Solution:

Let us organize this into parts.

a. Prove the following result:

$$
\int d^4 k \ \delta(k^\mu k_\mu - m^2)\theta(k^0) f(k^0, \vec{k}) = \int d^3 k \ \frac{1}{2\omega_k} f(\omega_k, \vec{k}) \ , \quad \omega_k \equiv \sqrt{|\vec{k}|^2 + m^2}
$$

b. Argue that $d^3 k/(2\omega_k)$ is a Lorentz invariant measure.
c. Work out $[a_k, a_{k'}]$, $[a_k^\dagger, a_{k'}^\dagger]$ and $[a_k, a_{k'}^\dagger]$ if the Fourier expansion for φ is:

$$
\varphi(\vec{x}, t) = \int \frac{d^D k}{(2\pi)^D 2\omega_k} \left[a_{\vec{k}} e^{-i\omega_k t + i\vec{k}\cdot\vec{x}} + a_{\vec{k}}^\dagger e^{i\omega_k t - i\vec{k}\cdot\vec{x}} \right]
$$

Solution:

a. This follows almost immediately from the following property of the delta function:

$$
\delta\left(F(x)\right) = \sum_{x_*} \frac{1}{|F'(x_*)|} \delta(x - x_*) \ , \quad \text{where } F(x_*) \equiv 0
$$

For this problem, define $F(k^0) \equiv (k^0)^2 - \omega_k^2$. The zeros of this function occur at $k^0 = \pm\omega_k$, and the derivative of the function evaluated at these zeros is:

$$
F'(k^0) = 2k^0 = \pm 2\omega_k
$$

Therefore,

$$\delta\left(F(k^0)\right) = \frac{1}{2\omega_k}\left(\delta(k^0 - \omega_k) + \delta(k^0 + \omega_k)\right)$$

We may now prove the desired result:

$$\int d^4k\,\delta(k^\mu k_\mu - m^2)\theta(k^0)f(k^0,\vec{k}) = \int d^3k\int_{-\infty}^{\infty} dk^0\,\delta((k^0)^2 - |\vec{k}|^2 - m^2)\theta(k^0)f(k^0,\vec{k})$$

$$= \int d^3k\int_{-\infty}^{\infty} dk^0\,\delta(F(k^0))\theta(k^0)f(k^0,\vec{k})$$

$$= \int d^3k\int_{-\infty}^{\infty} dk^0\left[\frac{1}{2\omega_k}\left(\delta(k^0 - \omega_k) + \delta(k^0 + \omega_k)\right)\right]\underbrace{\theta(k^0)}_{\text{picks out } k^0 > 0 \text{ only}}f(k^0,\vec{k})$$

$$= \int d^3k\int_{-\infty}^{\infty} dk^0\left[\frac{1}{2\omega_k}\delta(k^0 - \omega_k)\right]f(k^0,\vec{k})$$

$$= \int d^3k\,\frac{1}{2\omega_k}f(\omega_k,\vec{k})$$

b. Consider Lorentz transformations that do not change the sign of the time component of a vector. The matrices $\Lambda \equiv \partial x/\partial x'$ have determinant $+1$, so $\int d^4k\,\delta(k^2 - m^2)\theta(k^0)$ is unchanged under a Lorentz transformation. The function $f(k)$ has no Lorentz indices and therefore does not transform under Lorentz transformations. Therefore the whole integral $\int d^4k\,\delta(k^2 - m^2)\theta(k^0)f(k)$ is Lorentz invariant.

If the integral $\int d^3k\frac{1}{2\omega_k}f(\omega_k,\vec{k})$ is to be Lorentz invariant, and if the function $f(\omega_k,\vec{k})$ is Lorentz invariant, then the only conclusion we can make is that the measure $d^3k/(2\omega_k)$ is also Lorentz invariant.

c. The point is to find the new commutation relations if you do not keep the square root in the definition of the Fourier coefficients $a_{\vec{k}}$ and $a^\dagger_{\vec{k}}$. Assume the relation $[a_k, a^\dagger_p] = \delta(k - p)$ then shift $a_k \to a_k/\sqrt{(2\pi)^D 2\omega_k}$ to get:

$$[a_{\vec{k}},\, a_{\vec{p}}] = (2\pi)^D 2\omega_{\vec{k}}\,\delta^{(D)}(\vec{k} - \vec{p})$$

If you like, we can compute this explicitly as well. Using the square-root integration measure, the Fourier expansion of $\varphi(\vec{x},t)$ is:

$$\varphi(\vec{x},t) = \int \frac{d^D k}{\sqrt{(2\pi)^D\,2\omega_k}}\left[a_{\vec{k}}\,e^{-i(\omega_k t - \vec{k}\cdot\vec{x})} + a^\dagger_{\vec{k}}\,e^{i(\omega_k t - \vec{k}\cdot\vec{x})}\right]$$

The canonical commutation relations are between φ and $\pi \equiv \partial\mathcal{L}/\partial\dot{\varphi} = \dot{\varphi}$, so we also need the Fourier expansion of π:

22

$$\pi(\vec{x}, t) = -i \int \frac{d^D k}{\sqrt{(2\pi)^D}} \sqrt{\frac{\omega_k}{2}} \left[a_{\vec{k}} e^{-i(\omega_k t - \vec{k} \cdot \vec{x})} - a_{\vec{k}}^\dagger e^{i(\omega_k t - \vec{k} \cdot \vec{x})} \right]$$

Now that we've taken the time derivative we may as well set $t = 0$ to simplify the calculation. We need to change basis from (φ, π) to $(a, a^\dagger)$, and to do that we can invert the Fourier transform:

$$\int d^D x \, e^{-i\vec{p} \cdot \vec{x}} \, \varphi(\vec{x}, 0) = \int \frac{d^D k}{\sqrt{(2\pi)^D 2\omega_k}} \int d^D x \left[a_{\vec{k}} e^{i(\vec{k} - \vec{p}) \cdot \vec{x}} + a_{\vec{k}}^\dagger e^{-i(\vec{k} + \vec{p}) \cdot \vec{x}} \right]$$

$$= \int d^D k \sqrt{\frac{(2\pi)^D}{2\omega_k}} \left[a_{\vec{k}} \delta^{(D)}(\vec{k} - \vec{p}) + a_{\vec{k}}^\dagger \delta^{(D)}(\vec{k} + \vec{p}) \right]$$

$$= \sqrt{\frac{(2\pi)^D}{2\omega_p}} \left(a_{\vec{p}} + a_{-\vec{p}}^\dagger \right)$$

Note that $\omega_p = \sqrt{|\vec{p}|^2 + m^2}$, so the sign of $\vec{p}$ does not affect ω_p. The algebra goes through in exactly the same way for π except with the relative minus sign:

$$\int d^D x \, e^{-i\vec{p} \cdot \vec{x}} \pi(\vec{x}, 0) = -i \sqrt{\frac{(2\pi)^D \omega_p}{2}} \left(a_{\vec{p}} - a_{-\vec{p}}^\dagger \right)$$

Rearranging these for clarity, we have:

$$a_{\vec{p}} + a_{-\vec{p}}^\dagger = \sqrt{\frac{2\omega_p}{(2\pi)^D}} \int d^D x \, e^{-i\vec{p} \cdot \vec{x}} \, \varphi(\vec{x}, 0)$$

$$a_{\vec{p}} - a_{-\vec{p}}^\dagger = i \sqrt{\frac{2}{(2\pi)^D \omega_p}} \int d^D x \, e^{-i\vec{p} \cdot \vec{x}} \pi(\vec{x}, 0)$$

Organized in this way it's pretty clear that what we want to do is to add the two equations and divide by 2. Once we have $a_{\vec{p}}$ we can just Hermitian conjugate it to get $a_{\vec{p}}^\dagger$. Therefore:

$$a_{\vec{p}} = \frac{1}{\sqrt{(2\pi)^D 2\omega_p}} \int d^D x \, e^{-i\vec{p} \cdot \vec{x}} \left[\omega_p \varphi(\vec{x}, 0) + i\pi(\vec{x}, 0) \right]$$

The non-zero equal-time canonical commutation relation is $[\varphi(\vec{x}, t), \pi(\vec{y}, t)] = i\delta^{(D)}(\vec{x} - \vec{y})$. The other commutators are zero. Therefore we can compute $[a_{\vec{p}}, a_{\vec{k}}]$:

$$[a_{\vec{p}}, a_{\vec{k}}] = \frac{1}{(2\pi)^D 2\sqrt{\omega_p \omega_k}} \int d^D x \, e^{-i\vec{p}\cdot\vec{x}} \int d^D y \, e^{-i\vec{k}\cdot\vec{y}} [\,\omega_p \varphi(\vec{x}) + i\pi(\vec{x})\,,\, \omega_k \varphi(\vec{y}) + i\pi(\vec{y})\,]$$

$$= \frac{1}{(2\pi)^D 2\sqrt{\omega_p \omega_k}} \int d^D x \, e^{-i\vec{p}\cdot\vec{x}} \int d^D y \, e^{-i\vec{k}\cdot\vec{y}} \left(0 + i\omega_p \underbrace{[\varphi(\vec{x}), \pi(\vec{y})]}_{i\,\delta^{(D)}(\vec{x}-\vec{y})} + i\omega_k \underbrace{[\pi(\vec{x}), \varphi(\vec{y})]}_{-i\,\delta^{(D)}(\vec{x}-\vec{y})} + 0 \right)$$

$$= \frac{1}{(2\pi)^D 2\sqrt{\omega_p \omega_k}} \int d^D x \, e^{-i(\vec{p}+\vec{k})\cdot\vec{x}} \left(-\omega_p + \omega_k \right)$$

$$= \frac{1}{2\sqrt{\omega_p \omega_k}} \delta^{(D)}(\vec{k}+\vec{p}) \left(-\omega_p + \omega_k \right) = 0 \quad \text{since} \quad \omega_{-\vec{p}} = \omega_{+\vec{p}}$$

To get a non-zero result, we're going to have to change the sign on one term but not the other; but that's exactly what Hermitian conjugation will do for us. So, without spending much extra effort we can change the sign in the e^{ipx} factor and the sign of ω_p to get:

$$[a_{\vec{p}}, a_{\vec{k}}^\dagger] = \frac{1}{2\sqrt{\omega_p \omega_k}} \delta^{(D)}(\vec{k}-\vec{p})(+\omega_p + \omega_k) = \delta^{(D)}(\vec{k}-\vec{p})$$

Finally, $[a_{\vec{p}}^\dagger, a_{\vec{k}}^\dagger] = 0$ by Hermitian conjugating $[a_{\vec{p}}, a_{\vec{k}}] = 0$.

3. For the complex scalar field discussed in the text calculate $\langle 0|T[\varphi(x)\varphi^\dagger(0)]|0\rangle$.

Solution:

We want to calculate $\langle 0|T[\varphi(\vec{x},t)\varphi(\vec{0},0)]|0\rangle$, $\langle 0|T[\varphi(\vec{x},t)\varphi^\dagger(\vec{0},0)]|0\rangle$ and $\langle 0|T[\varphi^\dagger(\vec{x},t)\varphi^\dagger(\vec{0},0)]|0\rangle$. For convenience, let's write the Fourier expansions for φ and $\varphi^\dagger$ again:

$$\varphi(\vec{x},t) = \int \frac{d^3 k}{(2\pi)^3 2\omega_{\vec{k}}} \left(a_{\vec{k}} \, e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{x})} + b_{\vec{k}}^\dagger \, e^{i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{x})} \right)$$

$$\varphi^\dagger(\vec{x},t) = \int \frac{d^3 k}{(2\pi)^3 2\omega_{\vec{k}}} \left(a_{\vec{k}}^\dagger \, e^{+i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{x})} + b_{\vec{k}} \, e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{x})} \right)$$

The vacuum state is defined by $a_{\vec{k}}|0\rangle = b_{\vec{k}}|0\rangle = 0$. Since a commutes with $b^\dagger$, we see immediately that the two-point function of φ with itself is zero. Similarly, the two-point function of $\varphi^\dagger$ with itself is zero. So we just have to calculate the two point function for φ with its Hermitian conjugate. For $t > 0$, we have:

$$\langle 0|\varphi(\vec{x},t)\varphi^\dagger(\vec{0},0)|0\rangle = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \int \frac{d^3p}{(2\pi)^3 2\omega_p} \langle 0| \left(a_{\vec{k}}\, e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{x})} + b_{\vec{k}}^\dagger\, e^{i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{x})} \right) \left(a_{\vec{p}}^\dagger + b_{\vec{p}} \right) |0\rangle$$

$$= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \int \frac{d^3p}{(2\pi)^3 2\omega_p}\, e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{x})} \quad \underbrace{\langle 0|a_{\vec{k}}\, a_{\vec{p}}^\dagger|0\rangle}_{(2\pi)^3 2\omega_p\, \delta^{(3)}(\vec{k} - \vec{p})}$$

$$= \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, e^{-i(\omega_k t - \vec{k}\cdot\vec{x})}$$

We also have, for the opposite ordering:

$$\langle 0|\varphi^\dagger(\vec{0},0)\varphi(\vec{x},t)|0\rangle = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \int \frac{d^3p}{(2\pi)^3 2\omega_p} \langle 0| \left(a_{\vec{p}}^\dagger + b_{\vec{p}} \right) \left(a_{\vec{k}}\, e^{-i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{x})} + b_{\vec{k}}^\dagger\, e^{i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{x})} \right) |0\rangle$$

$$= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \int \frac{d^3p}{(2\pi)^3 2\omega_p}\, e^{i(\omega_{\vec{k}}t - \vec{k}\cdot\vec{x})} \quad \underbrace{\langle 0|b_{\vec{p}}\, b_{\vec{k}}^\dagger|0\rangle}_{(2\pi)^3 2\omega_p\, \delta^{(3)}(\vec{k} - \vec{p})}$$

$$= \int \frac{d^3k}{(2\pi)^3 2\omega_k}\, e^{i(\omega_k t - \vec{k}\cdot\vec{x})}$$

So the time ordered product of the two fields is:

$$\langle 0|T[\varphi(\vec{x},t)\varphi^\dagger(\vec{0},0)]|0\rangle \equiv \theta(t)\langle 0|\varphi(\vec{x},t)\varphi^\dagger(\vec{0},0)|0\rangle + \theta(-t)\langle 0|\varphi^\dagger(\vec{0},0)\varphi(\vec{x},t)|0\rangle$$

$$= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \left(\theta(t)e^{-i(\omega_k t - \vec{k}\cdot\vec{x})} + \theta(-t)e^{i(\omega_k t - \vec{k}\cdot\vec{x})} \right)$$

$$= i \int \frac{d^4k}{(2\pi)^4} \frac{e^{ik_\mu x^\mu}}{k^2 - m^2 + i\varepsilon}$$

I.9 Disturbing the Vacuum

1. Choose the damping function $g(v) = 1/(1+v)^2$ instead of the one in the text. Show that this results in the same Casimir force. [Hint: To sum the resulting series, pass to an integral representation $H(\xi) = -\sum_{n=1}^{\infty} 1/(1+n\xi) = -\sum_{n=1}^{\infty} \int_0^{\infty} dt\, e^{-(1+n\xi)t} = \int_0^{\infty} dt\, e^{-t}/(1-e^{\xi t})$. Note that the integral blows up logarithmically near the lower limit, as expected.]

Solution:

The function $f(d)$ is

$$f(d) = \frac{\pi}{2d} \sum_{n=1}^{\infty} n\, g\left(\frac{n\pi}{d}a\right) = \frac{1}{2}\frac{\partial}{\partial a} \sum_{n=1}^{\infty} h\left(\frac{n\pi}{d}a\right)$$

If $g(v) = (1+v)^{-2}$, then $h(v) = \int dv\, g(v) = -(1+v)^{-1}$ up to some constant that we can set equal to zero. Define $\varepsilon \equiv \pi a/d$. Using the hint, we can rewrite the function $f(d)$ as an integral:

$$\begin{aligned}
f(d) &= \frac{1}{2}\frac{\partial}{\partial a} \sum_{n=1}^{\infty} h\left(n\varepsilon\right) = -\frac{1}{2}\frac{\partial}{\partial a} \sum_{n=1}^{\infty} \frac{1}{1+n\varepsilon} \\
&= -\frac{1}{2}\frac{\partial}{\partial a} \sum_{n=1}^{\infty} \int_0^{\infty} dt\, e^{-(1+n\varepsilon)t} \\
&= -\frac{1}{2}\frac{\partial}{\partial a} \int_0^{\infty} dt\, e^{-t} \sum_{n=1}^{\infty} \left(e^{-\varepsilon t}\right)^n \\
&= -\frac{1}{2}\frac{\partial}{\partial a} \int_0^{\infty} dt\, \frac{e^{-t}}{1-e^{-\varepsilon t}}
\end{aligned}$$

Now we can Taylor expand the denominator, being careful to keep higher order terms:

$$e^{-\varepsilon t} = 1 - \varepsilon t + \frac{1}{2}(\varepsilon t)^2 - \frac{1}{6}(\varepsilon t)^3 + O(\varepsilon^4)$$

So, the denominator is:

$$\begin{aligned}
1 - e^{-\varepsilon t} &= +\varepsilon t - \frac{1}{2}(\varepsilon t)^2 + \frac{1}{6}(\varepsilon t)^3 + O(\varepsilon^4) \\
&= \varepsilon t \left[1 - \frac{1}{2}\varepsilon t + \frac{1}{6}(\varepsilon t)^2 + O(\varepsilon^3)\right]
\end{aligned}$$

We'd like to get that series out of the denominator, so use the expansion $(1+\delta)^{-1} = 1 - \delta +$

26

$\delta^2 + O(\delta^3)$ with $\delta = -\frac{1}{2}\varepsilon t + \frac{1}{6}(\varepsilon t)^2 + O(\varepsilon^3)$:

$$f(d) = -\frac{1}{2}\frac{\partial}{\partial a}\int_0^\infty dt\,\frac{e^{-t}}{\varepsilon t}\left[1 - \left(-\frac{1}{2}\varepsilon t + \frac{1}{6}(\varepsilon t)^2 + O(\varepsilon^3)\right) + \left(-\frac{1}{2}\varepsilon t + O(\varepsilon^2)\right)^2 + O(\varepsilon^3)\right]$$

$$= -\frac{1}{2}\frac{\partial}{\partial a}\int_0^\infty dt\,\frac{e^{-t}}{\varepsilon t}\left[1 + \frac{1}{2}\varepsilon t + \frac{1}{12}(\varepsilon t)^2 + O(\varepsilon^3)\right]$$

$$= -\frac{1}{2}\frac{\partial}{\partial a}\int_0^\infty dt\,e^{-t}\left[\frac{1}{\varepsilon t} + \frac{1}{2} + \frac{1}{12}\varepsilon t + O(\varepsilon^2)\right]$$

$$= -\frac{1}{2}\frac{\partial}{\partial a}\int_0^\infty dt\,e^{-t}\left[\frac{d}{\pi at} + \frac{1}{2} + \frac{\pi at}{12\,d} + O(a^2)\right]$$

$$= -\frac{1}{2}\int_0^\infty dt\,e^{-t}\left[-\frac{d}{\pi a^2 t} + 0 + \frac{\pi t}{12\,d} + O(a)\right]$$

Notice that differentiating with respect to a kills the second term and turns the $O(a)$ term into a finite piece; that is why we needed to go to third order in the original expansion. There is one integral we have to do: $\int_0^\infty dt\,e^{-t}\,t = \int_0^\infty dt\,e^{-t} = 1$. The other integral is just some number that will cancel out, so call it $I \equiv \int_0^\infty dt\,e^{-t}/t$ and proceed:

$$f(d) = +\frac{Id}{2\pi a^2} - \frac{\pi}{24\,d}$$

We have taken the limit $a \to 0$ to kill the finite a-dependent terms. Now we're in business. Differentiate with respect to d to get:

$$f'(d) = \frac{I}{2\pi a^2} + \frac{\pi}{24\,d^2}$$

Just as in the book for the case of the exponential damping factor, the infinite term is independent of the distance d, so subtracting off $f'(L-d)$ removes this term and leaves only the finite contribution. Remembering that the force is minus the finite piece of the above derivative, we get the same Casimir force as before:

$$F(d) = -f'(d) + \lim_{L\to\infty} f'(L-d) = -\frac{\pi}{24\,d^2}$$

2. Show that with the regularization used in the appendix, the $1/d$ expansion of the force between two conducting plates contains only even powers.

Solution:

Using the regularization given in on p. 74, we have:

$$f(d) = -\frac{\pi}{2}\sum_\alpha \frac{c_\alpha}{b_\alpha}\int_0^\infty dt\,e^{-t}\left(\frac{1}{1 - e^{-b_\alpha t/d}} - 1\right)$$

27

where $b_\alpha = \frac{\pi}{\Lambda_\alpha} \to 0$ regulate the series. The constants c_α and b_α are subject to the constraints

$$\sum_\alpha \frac{c_\alpha}{b_\alpha} = 0 \;, \;\; \sum_\alpha \frac{c_\alpha}{b_\alpha^2} = 0 \;, \;\; \sum_\alpha c_\alpha = 1 \;.$$

The force is $F = -[f'(d) - (d \to L - d)]$ with $L \to \infty$, where in the limit $L \to \infty$ the second term serves simply to cancel off the leading divergent piece. The derivative of $f(d)$ is

$$f'(d) = -\frac{\pi}{2d^2} \sum_\alpha c_\alpha \int_0^\infty dt \; e^{-t} t \; \frac{e^{b_\alpha t/d}}{\left(e^{b_\alpha t/d} - 1\right)^2} \;.$$

The function

$$g(x) \equiv \frac{x^2 e^x}{(e^x - 1)^2} = \tfrac{1}{2} x^2 \operatorname{csch}^2 x$$

is even in x. Therefore its Taylor series contains only even powers of x. The function we are interested in is $x^{-2} g(x)$ with $x = b_\alpha t/d$, with the first term to be canceled by subtracting $f'(L - d)$. Therefore the series is even in $1/d$.

3. Show off your skill in doing integrals by calculating the Casimir force in $(3+1)$-dimensional spacetime. For help, see M. Kardar and R. Golestanian, *Rev. Mod. Phys.* 71: 1233,1999; J. Feinberg, A. Mann, and M. Revzen, *Ann. Phys.* 288:103, 2001.

Solution:

As given in the book, the energy per unit plate-area is given by adding up the energies of all the individual modes:

$$\varepsilon(d) = \sum_{k_x, k_y, k_z} \omega(k_x, k_y, k_z)$$

For a massless scalar field quantized in the x-direction, the dispersion relation is:

$$\omega(k_x, k_y, k_z) = \sqrt{|\vec{k}|^2} = \sqrt{\left(\frac{n\pi}{d}\right)^2 + k_y^2 + k_z^2}$$

Since k_y and k_z are continuous, the discrete sum indicated above should actually be an integral: $\sum_{k_y} \to \int \frac{dk_y}{2\pi}$, and the same for k_z. Therefore, the energy per unit area between the plates at $x = 0$ and $x = d$ is:

$$\varepsilon(d) = \sum_{n=1}^\infty \int_{-\infty}^\infty \frac{dk_y}{2\pi} \int_{-\infty}^\infty \frac{dk_z}{2\pi} \sqrt{\left(\frac{n\pi}{d}\right)^2 + k_y^2 + k_z^2}$$

As in the $(1+1)$-dimensional case, we've made a mistake in writing down this formula:

28

we haven't accounted for the important physical fact that we can't contain arbitrarily high frequencies within the plates. Regularizing with the function $F(\omega, a) = e^{-a\omega}$, the energy per unit area between the plates at $x = 0$ and $x = d$ is:

$$\varepsilon(d) = \lim_{a \to 0} \sum_{n=1}^{\infty} \int_{-\infty}^{\infty} \frac{dk_y}{2\pi} \int_{-\infty}^{\infty} \frac{dk_z}{2\pi} \sqrt{\left(\frac{n\pi}{d}\right)^2 + k_y^2 + k_z^2} \, e^{-a\sqrt{\left(\frac{n\pi}{d}\right)^2 + k_y^2 + k_z^2}}$$

First let's do the integrals. Since the integrand depends only on the combination $k^2 \equiv k_y^2 + k_z^2$, change to polar coordinates:

$$\varepsilon(d) = \lim_{a \to 0} \sum_{n=1}^{\infty} \underbrace{\int_0^{2\pi} \frac{d\phi}{2\pi}}_{=1} \int_0^{\infty} \frac{dk\, k}{2\pi} \sqrt{\left(\frac{n\pi}{d}\right)^2 + k^2} \, e^{-a\sqrt{\left(\frac{n\pi}{d}\right)^2 + k^2}}$$

Let $c \equiv n\pi/d$ for convenience, and while doing the integral over k don't forget that c is a function of n. Change integration variables to $x^2 \equiv c^2 + k^2 \implies k\, dk = x\, dx$:

$$\varepsilon(d) = \lim_{a \to 0} \sum_{n=1}^{\infty} \int_{c(n)}^{\infty} \frac{dx\, x}{2\pi} x\, e^{-a\, x} = \frac{1}{2\pi} \lim_{a \to 0} \sum_{n=1}^{\infty} \int_{c(n)}^{\infty} dx\, x^2\, e^{-a\, x}$$

The integral can be done via integration by parts twice:

$$\int_c^{\infty} dx\, x^2\, e^{-a\, x} = \frac{-1}{a} \left[x^2 e^{-a\, x} \big|_c^{\infty} - 2 \int_c^{\infty} dx\, x\, e^{-a\, x} \right]$$

$$= \frac{-1}{a} \left[-c^2 e^{-ac} + \frac{2}{a} \left(x\, e^{-a\, x} \big|_c^{\infty} - \int_c^{\infty} dx\, e^{-a\, x} \right) \right]$$

$$= \frac{-1}{a} \left[-c^2 e^{-ac} + \frac{2}{a} \left(-c\, e^{-ac} - \frac{1}{a} e^{-ac} \right) \right]$$

$$= \frac{+1}{a} \left[c^2 + \frac{2c}{a} + \frac{2}{a^2} \right] e^{-ac}$$

To summarize, the energy per unit area is now:

$$\varepsilon(d) = \frac{1}{2\pi} \lim_{a \to 0} \frac{1}{a} \sum_{n=1}^{\infty} \left(c(n)^2 + \frac{2c(n)}{a} + \frac{2}{a^2} \right) e^{-ac(n)}, \text{ where } c(n) \equiv \frac{n\pi}{d}$$

The parameter a has dimensions of length, so to Taylor expand we should define a dimensionless ratio, $\alpha \equiv a/D$, where $D \equiv d/\pi$. The π was included in the α purely for notational

convenience. In terms of these variables, $c = n/D$ and $a = \alpha D$, so the energy is:

$$\varepsilon(d) = \frac{1}{2\pi D^3} \lim_{\alpha \to 0} \frac{1}{\alpha} \sum_{n=1}^{\infty} \left(n^2 + \frac{2n}{\alpha} + \frac{2}{\alpha^2} \right) e^{-\alpha n}$$

$$= \frac{1}{2\pi(d/\pi)^3} \lim_{\alpha \to 0} \frac{1}{\alpha} \left(\partial_\alpha^2 - \frac{2}{\alpha} \partial_\alpha + \frac{2}{\alpha^2} \right) \sum_{n=1}^{\infty} e^{-\alpha n}$$

$$= \frac{\pi^2}{d^3} \lim_{\alpha \to 0} \frac{1}{\alpha} \left(\frac{1}{2} \partial_\alpha^2 - \frac{1}{\alpha} \partial_\alpha + \frac{1}{\alpha^2} \right) \sum_{n=1}^{\infty} e^{-\alpha n}$$

The sum is an infinite geometric series:

$$\sum_{n=1}^{\infty} e^{-\alpha n} = \sum_{n=1}^{\infty} (e^{-\alpha})^n = \sum_{n=0}^{\infty} (e^{-\alpha})^n - 1$$

$$= \frac{1}{1 - e^{-\alpha}} - 1 = \frac{e^{-\alpha}}{1 - e^{-\alpha}} = \frac{1}{e^{\alpha} - 1}$$

Therefore, the energy is:

$$\varepsilon(d) = \frac{\pi^2}{d^3} \lim_{\alpha \to 0} \frac{1}{\alpha} \left(\frac{1}{2} \partial_\alpha^2 - \frac{1}{\alpha} \partial_\alpha + \frac{1}{\alpha^2} \right) \left(\frac{1}{e^{\alpha} - 1} \right)$$

The tedious partial derivatives are:

$$\partial_\alpha \left(\frac{1}{e^{\alpha} - 1} \right) = \frac{-e^{\alpha}}{(e^{\alpha} - 1)^2}$$

$$\partial_\alpha^2 \left(\frac{1}{e^{\alpha} - 1} \right) = \frac{e^{2\alpha} + e^{\alpha}}{(e^{\alpha} - 1)^3}$$

The energy is now:

$$\varepsilon(d) = \frac{\pi^2}{d^3} \lim_{\alpha \to 0} \frac{1}{\alpha} \left[\frac{e^{2\alpha} + e^{\alpha}}{2(e^{\alpha} - 1)^3} + \frac{e^{\alpha}}{\alpha(e^{\alpha} - 1)^2} + \frac{1}{\alpha^2(e^{\alpha} - 1)} \right]$$

The denominators go to zero as $\alpha \to 0$, but hidden within the divergent expression is a finite, physically meaningful result. To see that there is in fact a finite piece hidden in there, compute the Taylor series in α about $\alpha = 0$ for the energy (don't forget the $1/\alpha$ prefactor). The answer is:

$$\varepsilon(d) = \lim_{\alpha \to 0} \left[\underbrace{\frac{\pi^2}{d^3} \left(\frac{3}{\alpha^4} - \frac{1}{2\alpha^3} \right)}_{} \underbrace{-\frac{\pi^2}{720\, d^3}}_{\text{independent of } \alpha} + \underbrace{O(\alpha^2)}_{= 0 \text{ when } \alpha \to 0} \right]$$

30

Recalling the definition $\alpha = a/d$ gives:

$$\varepsilon(d) = \lim_{a \to 0} \frac{\pi^2}{d^3} \left(\frac{3\,d^4}{a^4} - \frac{d^3}{2\,a^3} \right) - \frac{\pi^2}{720\,d^3}$$

$$= \lim_{a \to 0} \pi^2 \left(\frac{3\,d}{a^4} - \frac{1}{2\,a^3} \right) - \frac{\pi^2}{720\,d^3}$$

The force per unit area between the pairs of plates at $x = 0$ and $x = d$ and at $x = d$ and $x = L$ as a function of the location $x = d$ of the middle plate is:

$$f(d) = -\frac{\partial\,\varepsilon_T(d)}{\partial d} = -\frac{\partial\,\varepsilon(d)}{\partial d} - \frac{\partial\,\varepsilon(L - d)}{\partial d}$$

$$= -\frac{\partial\,\varepsilon(d)}{\partial d} - \frac{\partial\,\varepsilon(L - d)}{\partial(L - d)} \frac{\partial(L - d)}{\partial d}$$

$$= -\frac{\partial\,\varepsilon(d)}{\partial d} + \frac{\partial\,\varepsilon(d)}{\partial d} \bigg|_{d \to L - d}$$

The derivative of ε with respect to the distance d is:

$$\frac{\partial\,\varepsilon(d)}{\partial d} = \lim_{a \to 0} \left(\frac{3\pi^2}{a^4} \right) + \frac{\pi^4}{240\,d^4}$$

The divergent piece is independent of the distance d; when we subtract the contribution to the force from the plate at $x = L$, the infinity arising from $\lim_{a \to 0} 1/a^4$ will cancel out:

$$\frac{\partial\,\varepsilon(d)}{\partial d} \bigg|_{d \to L - d} = \lim_{a \to 0} \left(\frac{3\pi^2}{a^4} \right) + \frac{\pi^2}{240\,(L - d)^4}$$

$$= \lim_{a \to 0} \left(\frac{3\pi^2}{a^4} \right) + \frac{\pi^2}{240\,L^4} \left[1 + O\left(\frac{d}{L} \right) \right]$$

$$\to \lim_{a \to 0} \left(\frac{3\pi^2}{a^4} \right) \quad \text{for } L \to \infty$$

The force per unit area between two parallel conducting plates a distance d apart is:

$$f(d) = -\frac{\pi^2}{240\,d^4}$$

The force is attractive, meaning that it takes effort to pull apart two conducting plates. Restoring the factors of $\hbar$ and c to give an idea for how large this force is gives:

$$f(d) = -\frac{\pi^2 \hbar c}{240\,d^4} \approx -1.3 \times 10^{-27} \left(\frac{\text{m}}{d} \right)^4 \text{N} \cdot \text{m}^{-2}$$

So the force (per unit area) between plates of size ~ 1 m^2 separated a distance $d \sim 1$ mm $= 10^{-3}$ m apart is $\sim 10^{-15}$ N. Since 1 N is about the gravitational force exerted by the Earth on an apple, the Casimir force is extremely small, as expected from the fact that it is an effect of quantum origin.

Note: Each harmonic oscillator mode contributions $\frac{1}{2}\omega$ to the vacuum energy, so for a real scalar field, which has only one component, this answer should be divided by 2. We keep it as is because it is the usual result quoted for the electromagnetic field, which has 2 polarizations. Also, if we decide to use a complex scalar field instead of a real scalar field, then we have a theory with two real scalar fields and for that case the answer would be correct as written.

I.10 Symmetry

1. Some authors prefer the following more elaborate formulation of Noether's theorem. Suppose that the action does not change under an infinitesimal transformation $\delta\varphi_a(x) = \theta^A V_a^A$ [with θ^A some parameters labeled by A and V_a^A some function of the fields $\varphi_b(x)$ and possibly also of their first derivatives with respect to x.] It is important to emphasize that when we say the action S does not change we are not allowed to use the equations of motion. After all, the Euler-Lagrange equations of motion follow from demanding that $\delta S = 0$ for any variation $\delta\varphi_a$ subject to certain boundary conditions. Our scalar field theory example nicely illustrates this point, which is confused in some books: $\delta S = 0$ merely because S is constructed using the scalar product of $O(N)$ vectors.

Now let us do something apparently a bit strange. Let us consider the infinitesimal change written above but with the parameters θ^A dependent on x. In other words, we now consider $\delta\varphi_a(x) = \theta^A(x)V_a^A$. Then of course there is no reason for δS to vanish; but, on the other hand, we know that since δS does vanish when θ^A is constant, δS must have the form $\delta S = \int d^4x\, J^\mu(x)\partial_\mu\theta^A(x)$. In practice, this gives us a quick way of reading off the current $J^\mu(x)$; it is just the coefficient of $\partial_\mu\theta^A(x)$ in δS.

Show how all this works for the Lagrangian in (3).

$$(3) \qquad \mathcal{L} = \frac{1}{2}\left[(\partial\vec{\varphi})^2 - m^2\vec{\varphi}^2\right] - \frac{1}{4}\lambda(\vec{\varphi}^2)^2$$

Solution:

Consider the following real scalar field Lagrangian with global $O(N)$ symmetry:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi_\alpha\partial^\mu\varphi^\alpha - \frac{1}{2}m^2\varphi_\alpha\varphi^\alpha - \frac{1}{4}\lambda(\varphi_\alpha\varphi^\alpha)$$

The field index α runs from 1 to N. The global $O(N)$ symmetry is:

$$O(N) : \varphi_\alpha \to R_\alpha{}^\beta\varphi_\beta \text{ with } R^T R = 1$$

Let $\{T^A\}_{A=1}^{N(N-1)/2}$ be the generators of the Lie algebra of the group $O(N)$ in the N-dimensional ("defining") representation. Given these generators, the $N \times N$ orthogonal matrix R can be written as:

$$R_\alpha{}^\beta = \left[e^{\theta^A T^A}\right]_\alpha{}^\beta = \delta_\alpha{}^\beta + \theta^A (T^A)_\alpha{}^\beta + O(\theta^2)$$

The goal is to find the conserved current corresponding to the $O(N)$ symmetry by promoting the symmetry to a local one and finding the coefficient of $\partial_\mu \theta^A$ in the variation of the action. Afterwards, we'll compare the result of this method to the result of the formal procedure for generating the Noether current and see that they are equivalent. The action is:

$$S[\varphi] = \int d^4x \left(\frac{1}{2}\partial_\mu\varphi_\alpha\partial^\mu\varphi^\alpha - \frac{1}{2}m^2\varphi_\alpha\varphi^\alpha - \frac{1}{4}\lambda(\varphi_\alpha\varphi^\alpha)\right)$$

Consider a first-order change in the fields:

$S[\varphi + \delta\varphi] =$
$$\int d^4x \left[\frac{1}{2}\partial_\mu(\varphi + \delta\varphi)_\alpha\partial^\mu(\varphi + \delta\varphi)^\alpha - \frac{1}{2}m^2(\varphi + \delta\varphi)_\alpha(\varphi + \delta\varphi)^\alpha - \frac{1}{4}\lambda((\varphi + \delta\varphi)_\alpha(\varphi + \delta\varphi)^\alpha)\right]$$
$$= S[\varphi] + \int d^4x \left(\partial_\mu\varphi^\alpha\partial^\mu(\delta\varphi_\alpha) - m^2\varphi^\alpha\,\delta\varphi_\alpha - \lambda\varphi^\beta\varphi_\beta\varphi^\alpha\,\delta\varphi_\alpha\right)$$

The $O(N)$ symmetry transformation is:

$$\varphi_\alpha \to R_\alpha{}^\beta\varphi_\beta = \varphi_\alpha + \theta^A(T^A)_\alpha{}^\beta\varphi_\beta + O(\theta^2)$$

Therefore, the corresponding first-order variation in the field is:

$$\delta\varphi_\alpha = \theta^A(T^A)_\alpha{}^\beta\varphi_\beta$$

The symmetry is global, but we can find the conserved current by imagining that the symmetry were local and finding a term of the form:

$$\delta S = \int d^4x\, j^\mu\partial_\mu\theta$$

If we use the explicit expression for $\delta\varphi$ in the first-order change in the action, we get:

$$\delta S \equiv S[\varphi + \delta\varphi] - S[\varphi] = \int d^4x \, \partial_\mu \varphi^\alpha \partial^\mu (\delta\varphi_\alpha) + \ldots$$

$$= \int d^4x \, \partial_\mu \varphi^\alpha \partial^\mu \left(\theta^A (T^A)_\alpha{}^\beta \varphi_\beta \right) + \ldots$$

$$= \int d^4x \, \partial_\mu \varphi^\alpha (T^A)_\alpha{}^\beta \varphi_\beta \left(\partial^\mu \theta^A \right)$$

We know the other terms must be zero because the action is invariant under global $O(N)$ transformations. Therefore, the conserved current corresponding to the global $O(N)$ symmetry is:

$$j_\mu^A = \partial_\mu \varphi^\alpha (T^A)_\alpha{}^\beta \varphi_\beta$$

You might worry that the current doesn't look hermitian, since $(T^A)^T = -T^A$, but once you transpose everything you will also switch the order of the $\partial\varphi$ and the φ, which means you have to integrate by parts and therefore generate a second minus sign to get back to the original form. So j_μ^A is hermitian.

To verify the result using Noether's theorem, we compute the current directly:

$$\theta^A j_\mu^A = \frac{\partial \mathcal{L}}{\partial(\partial^\mu \varphi_\alpha)} \delta\varphi_\alpha = \partial_\mu \varphi^\alpha \delta\varphi_\alpha$$

$$= \partial_\mu \varphi^\alpha \theta^A (T^A)_\alpha{}^\beta \varphi_\beta$$

So there it is, the same answer as obtained previously:

$$j_\mu^A = \partial_\mu \varphi^\alpha (T^A)_\alpha{}^\beta \varphi_\beta$$

4. Add a Lorentz scalar field η transforming as a vector under $SO(3)$ to the Lagrangian in exercise I.10.3, maintaining $SO(3)$ invariance. Determine the Noether currents in this theory. Using the equations of motion, check that the currents are conserved.

Solution:

The Lagrangian is $\mathcal{L} = \mathcal{L}_\varphi + \mathcal{L}_\eta + \mathcal{L}_{\varphi\eta}$, where:

$$\mathcal{L}_\varphi = \frac{1}{2} \text{Tr} \left(\partial_\mu \varphi \, \partial^\mu \varphi - m^2 \, \varphi\varphi \right) - \lambda \left[\text{Tr}(\varphi\varphi) \right]^2$$

$$\mathcal{L}_\eta = \frac{1}{2} \left(\partial_\mu \eta^\alpha \, \partial^\mu \eta_\alpha - \mu^2 \, \eta^\alpha \eta_\alpha \right) - g(\eta^\alpha \eta_\alpha)^2$$

$$\mathcal{L}_{\varphi\eta} = \kappa \, \eta_\alpha \eta_\beta \varphi_{\alpha\beta} = \kappa \, \vec{\eta}^T \varphi \, \vec{\eta} \, .$$

34

The index α runs from 1 to 3, and the indices are contracted with usual identity matrix.[2] φ transforms under the 5-representation of $SO(3)$, which can either be written as a 5-dimensional column vector or as a 3×3 symmetric traceless matrix. For this problem, we choose the latter description and write:

$$\varphi = \varphi_{\alpha\beta} = \frac{1}{2}(\varphi_{\alpha\beta} + \varphi_{\beta\alpha}) - \frac{1}{3}(\varphi_{\gamma\delta}\delta^{\gamma\delta})\,\delta_{\alpha\beta} = \begin{pmatrix} A & C & D \\ C & B & E \\ D & E & -(A+B) \end{pmatrix}$$

Note that the term $\mathrm{Tr}(\varphi\varphi\varphi\varphi)$ is not included, because in this case it is proportional to $[\mathrm{Tr}(\varphi\varphi)]^2$. For further details about φ, see the solution to problem 9.3 in the book.

The $SO(3)$ transformation acts on φ and on η by multiplying each index by a 3×3 orthogonal matrix with determinant 1:

$$\varphi_{\alpha\beta} \to R_\alpha{}^\gamma R_\beta{}^\delta \varphi_{\gamma\delta}$$
$$\eta_\alpha \to R_\alpha{}^\beta \eta_\beta$$
$$R^T R = 1 \ , \ \det R = 1$$

To find the Noether current, we need an infinitesimal version of the above transformations. If T^A are the generators of the 3-dimensional representation of $SO(3)$, where A runs from 1 to $3(3-1)/2 = 3$, then the matrix R can be written:

$$R_\alpha{}^\beta = \left[e^{\theta^A T^A}\right]_\alpha{}^\beta = \left[1 + \theta^A T^A + O(\theta^2)\right]_\alpha{}^\beta = \delta_\alpha{}^\beta + \theta^A (T^A)_\alpha{}^\beta + O(\theta^2)$$

Therefore, the infinitesimal versions of the $SO(3)$ transformations on φ and η are:

$$\varphi_{\alpha\beta} \to \left[\delta_\alpha{}^\gamma + \theta^A (T^A)_\alpha{}^\gamma + O(\theta^2)\right]\left[\delta_\beta{}^\delta + \theta^B (T^B)_\beta{}^\delta + O(\theta^2)\right]\varphi_{\gamma\delta}$$
$$= \left[\delta_\alpha{}^\gamma\delta_\beta{}^\delta + \theta^A\left((T^A)_\alpha{}^\gamma\,\delta_\beta{}^\delta + \delta_\alpha{}^\gamma\,(T^A)_\beta{}^\delta\right) + O(\theta^2)\right]\varphi_{\gamma\delta}$$
$$= \varphi_{\alpha\beta} + \theta^A \underbrace{\left[(T^A)_\alpha{}^\gamma\,\varphi_{\gamma\beta} + (T^A)_\beta{}^\delta\,\varphi_{\alpha\delta}\right]}_{= \,\delta\varphi_{\alpha\beta}^A} + O(\theta^2)$$

$$\eta_\alpha \to \left[\delta_\alpha{}^\beta + \theta^A (T^A)_\alpha{}^\beta + O(\theta^2)\right]\eta_\beta$$
$$= \eta_\alpha + \theta^A \underbrace{(T^A)_\alpha{}^\beta \eta_\beta}_{=\,\delta\eta_\alpha^A} + O(\theta^2)$$

---

[2]In $SO(3)$ there is no distinction between upper and lower indices. For this problem we use them interchangeably.

The Noether current is:

$$j_\mu^A = \frac{\partial \mathcal{L}}{\partial(\partial^\mu \varphi_{\alpha\beta})} \delta\varphi_{\alpha\beta}^A + \frac{\partial \mathcal{L}}{\partial(\partial^\mu \eta_\alpha)} \delta\eta_\alpha^A$$

$$= \partial_\mu \varphi^{\alpha\beta}\, \delta\varphi_{\alpha\beta}^A + \partial_\mu \eta^\alpha\, \delta\eta_\alpha^A$$

$$= \partial_\mu \varphi^{\alpha\beta} \left[(T^A)_\alpha{}^\gamma\, \varphi_{\gamma\beta} + (T^A)_\beta{}^\delta\, \varphi_{\alpha\delta} \right] + \partial_\mu \eta^\alpha (T^A)_\alpha{}^\beta\, \eta_\beta$$

Now we have to check that this current is in fact conserved, or in other words that $\partial^\mu j_\mu^A = 0$. The equations of motion are:

$$\left[\partial^2 + m^2 + \lambda \operatorname{Tr}(\varphi\varphi) \right] \varphi_{\alpha\beta} + \kappa\, \eta_\alpha \eta_\beta = 0$$

$$\left\{ \left[\partial^2 + \mu^2 + g\, (\vec{\eta} \cdot \vec{\eta}) \right] \delta_{\alpha\beta} + 2\kappa\, \varphi_{\alpha\beta} \right\} \eta_\beta = 0$$

The important point is that $\partial^2 \varphi_{\alpha\beta} = a(\varphi)\, \varphi_{\alpha\beta} - \kappa\, \eta_\alpha \eta_\beta$ and $\partial^2 \eta_\alpha = b(\eta)\, \eta_\alpha - 2\kappa\, \varphi_{\alpha\beta}\eta_\beta$, where a and b are $SO(3)$-scalar functions of φ and η, respectively. Therefore, the 4-divergence of the current is:

$$\partial^\mu j_\mu^A = a(\varphi) \left[\varphi^{\alpha\beta}(T^A)_\alpha{}^\gamma \varphi_{\gamma\beta} + \varphi^{\alpha\beta}(T^A)_\beta{}^\delta \varphi_{\alpha\delta} \right]$$

$$+ b(\eta)\, \eta^\alpha (T^A)_\alpha{}^\beta \eta_\beta$$

$$- 2\kappa\, \eta_\alpha [(T^A)_{\alpha\beta}\varphi_{\beta\gamma} + \varphi_{\alpha\beta}(T^A)_{\beta\gamma}]\eta_\gamma$$

Each of the three terms is zero individually by symmetry. To see why, look at the first term:

$$\varphi^{\alpha\beta}(T^A)_\alpha{}^\gamma \varphi_{\gamma\beta} = (T^A)_\alpha{}^\gamma \varphi_{\gamma\beta}\varphi^{\alpha\beta}$$

$$= (T^A)_\alpha{}^\gamma \varphi_{\gamma\beta}\varphi^{\beta\alpha}$$

$$= (T^A)_\alpha{}^\gamma (\varphi\varphi)_\gamma{}^\alpha$$

Now the symmetry argument comes in: $\{T^A\}_{A=1}^3$ are the generators of 3-dimensional rotations and are therefore antisymmetric matrices. Meanwhile, $(\varphi\varphi)$ is a symmetric matrix, so $\operatorname{Tr}\left[(T^A)(\varphi\varphi) \right] = 0$. Similarly, $\eta^\alpha (T^A)_\alpha{}^\beta \eta_\beta = \operatorname{Tr}\left[(T^A)(\eta\eta) \right] = 0$.

For the third term, $\vec{\eta}^T (T^A\varphi + \varphi T^A)\vec{\eta}$, we have $\varphi = \varphi^T$ and $T^A = -(T^A)^T$. Therefore:

$$\vec{\eta}^T T^A \varphi\, \vec{\eta} = \eta_\alpha (T^A)_{\alpha\beta}\varphi_{\beta\gamma}\eta_\gamma = \eta_\gamma \varphi_{\beta\gamma}(T^A)_{\alpha\beta}\eta_\alpha = \eta_\gamma[+\varphi_{\gamma\beta}][-(T^A)_{\beta\alpha}]\eta_\alpha = -\vec{\eta}^T \varphi\, T^A \vec{\eta}\,.$$

Therefore, $\partial^\mu j_\mu^A = 0$. The current is conserved.

36

I.11 Field Theory in Curved Spacetime

1. Integrate by parts to obtain for the scalar field action

$$S = - \int d^4x \, \sqrt{-g} \, \frac{1}{2} \varphi \left(\frac{1}{\sqrt{-g}} \partial_\mu \sqrt{-g} \, g^{\mu\nu} \partial_\nu + m^2 \right) \varphi$$

and write the equation of motion for φ in curved spacetime. Discuss the propagator of the scalar field $D(x, y)$ (which is of course no longer translation invariant, i.e., it is no longer a function of $x - y$).

Solution:

The equation of motion that follows from the above action is

$$\frac{1}{\sqrt{-g}} \left[\partial_\mu \left(\sqrt{-g} \, g^{\mu\nu} \partial_\nu \right) + m^2 \right] \varphi = 0 \, .$$

We will specialize to the maximally symmetric de Sitter spacetime, which has scalar curvature $R = n(n-1)/a^2$ in n spacetime dimensions. Here a is the de Sitter length. We will follow B. Allen and T. Jacobson, "Vector Two-Point Functions in Maximally Symmetric Spaces," Commun. Math. Phys. 103, 669-692 (1986).

Although the scalar propagator $D(x, y)$ is no longer a function of $x - y$, it is still a function of x and y only through the geodesic distance between the points x and y:

$$\mu(x, y) = \int_0^1 d\lambda \left(g_{\mu\nu} \frac{dX^\mu(\lambda)}{d\lambda} \frac{dX^\nu(\lambda)}{d\lambda} \right)^{1/2}$$

where $X^\mu(0) = x^\mu$ and $X^\mu(1) = y^\mu$.

Therefore the defining differential equation for the free scalar propagator in maximally symmetric curved space can also be written as an ordinary differential equation in one variable, μ. Writing $D(x, y) = D(\mu(x, y))$, the equation of motion for φ implies

$$D''(\mu) + (n-1)A(\mu)D'(\mu) - m^2 D(\mu) = 0$$

for $x \neq y$. Here $' = \frac{d}{d\mu}$ and $A(\mu) = \frac{1}{a} \cot(\mu/a)$. Change variables to

$$z = \cos^2 \left(\frac{\mu}{2a} \right)$$

to put the differential equation in the form

$$\left\{ z(1-z) \frac{d^2}{dz^2} + [c - (a+b+1)z] \frac{d}{dz} - ab \right\} D(z) = 0$$

where

$$a = \tfrac{1}{2}\left[n - 1 + \sqrt{(n-1)^2 - (2ma)^2}\right]$$
$$b = \tfrac{1}{2}\left[n - 1 - \sqrt{(n-1)^2 - (2ma)^2}\right]$$
$$c = \tfrac{1}{2}n \ .$$

This is the defining equation of the hypergeometric function $_2F_1(a, b, c, z)$. Since $a + b + 1 - 2c = 0$, the differential equation is invariant under $z \to 1 - z$. Two independent particular solutions are therefore $_2F_1(a, b, c, z)$ and $_2F_1(a, b, c, 1 - z)$, and the general solution is a linear combination of these:

$$D(\mu) = C_{(1)} \, _2F_1(a, b, c, z(\mu)) + C_{(2)} \, _2F_1(a, b, c, 1 - z(\mu)) \ .$$

To fix the coefficients, we need to consider the $\mu \to 0$ and $\mu \to \infty$ behavior of $D(\mu)$, as well as the location of singular points and branch cuts. For further discussion, consult the reference.

II Dirac and the Spinor

II.1 The Dirac Equation

7. Show explicitly that (25) violates parity.

$$\mathcal{L} = G\,(\bar{\psi}_{1L}\gamma^{\mu}\psi_{2L})(\bar{\psi}_{3L}\gamma_{\mu}\psi_{4L}) \qquad\qquad (25)$$

Solution:

Parity acts on a Dirac spinor as $\psi \to i\gamma^0\psi$. Define $P_L \equiv \frac{1}{2}(1-\gamma^5)$ and $P_R \equiv \frac{1}{2}(1+\gamma^5)$ as usual. Then parity acts on a left-handed spinor as:

$$\psi_L \equiv P_L\psi \to P_L i\gamma^0\psi = i\gamma^0 P_R\psi = i\gamma^0\psi_R\ .$$

Thus parity transforms left-handed spinors into right-handed spinors, and so the Lagrangian in (25) violates parity.

11. Work out the Dirac equation in (1+1)-dimensional spacetime.

See VII.7.7.

12. Work out the Dirac equation in (2+1)-dimensional spacetime. Show that the apparently innocuous mass term violates parity and time reversal. [Hint: The three γ^{μ}s are just the three Pauli matrices with appropriate factors of i.]

Solution:

Consider the Dirac equation in (2+1) spacetime:

$$(i\slashed{\partial} - m)\psi = 0 \implies (i\gamma^0\partial_0 + i\gamma^1\partial_1 + i\gamma^2\partial_2 - m)\psi(x) = 0$$

Multiply the whole equation by γ^2 and anticommute it through to the right to get:

$$(-i\gamma^0\partial_0 - i\gamma^1\partial_1 + i\gamma^2\partial_2 - m)\gamma^2\psi(x) = 0$$

Divide by -1 to get:

$$(i\gamma^0\partial_0 + i\gamma^1\partial_1 - i\gamma^2\partial_2 + m)\gamma^2\psi(x) = 0$$

Defining the parity transformation $P : x \equiv (x^0, x^1, x^2) \to x' \equiv (x^0, x^1, -x^2)$ puts the above equation into the form:

$$(i\gamma^0\partial_{0'} + i\gamma^1\partial_{1'} + i\gamma^2\partial_{2'} + m)\gamma^2\psi(x) = 0$$

Therefore the spinor $\psi'(x') \equiv \gamma^2 \psi(x)$ satisfies the Dirac equation with the wrong sign for the mass term: $(i\partial\!\!\!/_{x'} + m)\psi'(x') = 0$.

At this point we should ask why the same argument doesn't hold in (3+1) dimensions. After all, if we parity transform using $P : x \equiv (x^0, x^1, x^2, x^3) \rightarrow x' \equiv (x^0, x^1, -x^2, x^3)$, we get the exact same result as above: $(i\partial\!\!\!/_{x'} + m)\psi'(x') = 0$. Where does the argument fail?

There are two answers. One answer is that in (odd+1) dimensions, the operation of flipping one spatial coordinate is related by a rotation to the operation of flipping all of the spatial coordinates, while in (even+1) dimensions that is not true. See the footnote on page 98 of the book for more details.

The second answer is that in (3+1) dimensions, we can multiply $(i\,\partial\!\!\!/_{x'} + m)\psi'(x') = 0$ on the left by the γ^5 matrix, which anticommutes with the other 4 gamma matrices, to get $(-i\,\partial\!\!\!/_{x'} + m)\gamma^5\psi'(x') = 0 \implies (i\partial\!\!\!/_{x'} - m)\gamma^5\psi'(x') = 0$, and the field redefinition $\psi \rightarrow \gamma^5\psi$ leaves the path integral unchanged. Therefore we conclude that the sign of the Dirac mass term does not matter in (3+1) dimensions.

However, this operation is not possible in (2+1) dimensions. As noted in the partial solution to this question in the back of the book, in (2+1) dimensions we can define the gamma matrices to be the Pauli matrices: $\gamma^0 = \sigma^3, \gamma^1 = i\sigma^2, \gamma^2 = -i\sigma^1$. Now try to define γ^5:

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2 = i\sigma^3(i\sigma^2)(-i\sigma^1) = i\sigma^3(-i\sigma^3) = I$$

In (2+1) dimensions, γ^5 is just the identity matrix. In other words, the Dirac spinor representation of $SO(2,1)$ is irreducible. Therefore, in (2+1) dimensions, the sign of the Dirac mass term does matter and therefore the mass term violates parity.

Time reversal is going to work the same way. Take the Dirac equation and multiply by γ^0, then anticommute γ^0 all the way to the right. This yields

$$(i\gamma^0\partial_0 - i\gamma^i\partial_i - m)\gamma^0\psi(x) = 0$$

Divide by -1 to get

$$(-i\gamma^0\partial_0 + i\gamma^i\partial_i + m)\gamma^0\psi(x) = 0$$

Define the time reversed coordinates $x' \equiv (-x^0, \vec{x})$ to write the above as

$$(+i\gamma^0\partial_{0'} + i\gamma^i\partial_{i'} + m)\gamma^0\psi(x) = 0$$

Therefore the spinor $\psi'(x') \equiv \gamma^0\psi(x)$ satisfies the Dirac equation with the wrong sign for the mass term. Again, in (2+1) dimensions the Dirac spinor is irreducible.

II.2 Quantizing the Dirac Field

2. Quantize the Dirac field in a box of volume V and show that the vacuum energy E_0 is indeed proportional to V. [Hint: The integral over momentum $\int d^3p$ is replaced by a sum over discrete values of the momentum.]

Solution:

This problem was worded incorrectly in the book, since the vacuum energy is $E_0 = (\#) \sum_p \omega_p$ with no factor of volume (as required by dimensional analysis). The intention is to place the Dirac field in a box and to calculate its contribution to the energy of the vacuum, and show that the factors of volume work themselves out. The solution proceeds according to the discussion on pp. 111 and 139, so we outline the salient steps:

- When quantizing in a box, momentum integrals are replaced with sums as:

$$\int \frac{d^3p}{(2\pi)^3} \to \frac{1}{V} \sum_{\vec{p}}$$

- Each harmonic oscillator for a Lorentz-scalar field contributes $+\frac{1}{2}\omega_{\vec{p}}$ to the vacuum energy.

- Each harmonic oscillator for a Lorentz-spinor field contributes $-\frac{1}{2}\omega_{\vec{p}}$ to the vacuum energy.

- The Dirac field contains two Lorentz-spinor fields, for example the electron and the positron, so the above gets multiplied by 2.

- Each Lorentz-spinor field contains two spin states, so the above gets multiplied by yet another factor of 2.

Therefore, a Dirac field contributes a total of $-2 \sum_{\vec{p}} \omega_{\vec{p}}$ to the energy of the vacuum.

II.3   Lorentz Group and Weyl Spinors

1. Show by explicit computation that $(\frac{1}{2}, \frac{1}{2})$ is indeed the Lorentz vector.

Solution:

It is sufficient to compute the infinitesimal transformation properties, but it is pedagogically instructive to show that it all works out for the finite group transformation.

Let χ transform under the $(\frac{1}{2}, \frac{1}{2})$ representation of the Lorentz group: $\chi_a{}^{\dot c}$, where $a \in \{1, 2\}$ denotes the two-dimensional representation of $SU(2)_L$ and $\dot c \in \{\dot 1, \dot 2\}$ denotes the two-dimensional representation of $SU(2)_R$.

The Lorentz group consists of the rotations, which are generated by $\vec J$, and of the boosts, which are generated by $\vec K$. Let A transform under the 2-dimensional representation of $SU(2)_L$, which means that A carries one $SU(2)_L$ 2-index: A_a.

The rotations and the boosts act on this object A as:

$$\text{Rotation through angle } \theta \text{ about axis } \hat n : A_a \rightarrow \left[e^{i\theta\, \hat n \cdot \vec J}\right]_a{}^b A_b = \left[e^{i\theta\, \hat n \cdot \left(\frac{\vec\sigma}{2}\right)}\right]_a{}^b A_b$$

$$\text{Boost through } \beta \equiv \mathrm{Arctanh}(v) \text{ about } \hat n : A_a \rightarrow \left[e^{i\beta\, \hat n \cdot \vec K}\right]_a{}^b = \left[e^{i\beta\, \hat n \cdot \left(\frac{\vec\sigma}{2i}\right)}\right]_a{}^b A_b$$

Since the Pauli matrices have the property $\{\sigma^i, \sigma^j\} = 2\delta^{ij}$, these transformations simplify:

$$
\begin{aligned}
e^{i\theta\, \hat n \cdot \left(\frac{\vec\sigma}{2}\right)} &= 1_{2\times 2} \cos\left(\frac{\theta}{2}\right) + i\, \hat n \cdot \vec\sigma \sin\left(\frac{\theta}{2}\right) \\
&= \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & 0 \\ 0 & \cos\left(\frac{\theta}{2}\right) \end{pmatrix} + i\left[n^x \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + n^y \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + n^z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right] \sin\left(\frac{\theta}{2}\right) \\
&= \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) + i\, n^z \sin\left(\frac{\theta}{2}\right) & i\, n^- \sin\left(\frac{\theta}{2}\right) \\ i\, n^+ \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) - i\, n^z \sin\left(\frac{\theta}{2}\right) \end{pmatrix} \quad, \quad n^\pm \equiv n^x \pm i\, n^y
\end{aligned}
$$

Using $\cos(-i\,x) = \cosh(x)$ and $i\sin(-i\,x) = \sinh(x)$, the boost transformation matrix can be obtained from the above with $\theta \rightarrow -i\beta$:

$$
e^{\beta\, \hat n \cdot \left(\frac{\vec\sigma}{2}\right)} = \begin{pmatrix} \cosh\left(\frac{\beta}{2}\right) + n^z \sinh\left(\frac{\beta}{2}\right) & n^- \sinh\left(\frac{\beta}{2}\right) \\ n^+ \sinh\left(\frac{\beta}{2}\right) & \cosh\left(\frac{\beta}{2}\right) - n^z \sinh\left(\frac{\beta}{2}\right) \end{pmatrix}
$$

As indicated on page 114 of the book, the transformations for $SU(2)_R$ are the same as those for $SU(2)_L$ except that the boosts get an extra minus sign from the replacement $\beta \rightarrow -\beta$. We can now compute explicitly how the four components of $\chi_a{}^{\dot c}$ transform under rotations and under boosts. First specialize to the case $\hat n = (1, 0, 0)$, for which the transformations on χ are:

Rotation through angle θ about $\hat{x}$-axis:

$$\chi_a{}^{\dot{c}} \to \left[e^{i\theta\,\hat{n}\cdot\vec{J}}\right]_a{}^b \left[e^{i\theta\,\hat{n}\cdot\vec{J}}\right]_{\dot{e}}{}^{\dot{c}} \chi_b{}^{\dot{e}}$$

$$= \begin{pmatrix} c & i\,s \\ i\,s & c \end{pmatrix}_a{}^b \begin{pmatrix} c & i\,s \\ i\,s & c \end{pmatrix}_{\dot{e}}{}^{\dot{c}} \chi_b{}^{\dot{e}}$$

$$= \begin{pmatrix} c & i\,s \\ i\,s & c \end{pmatrix}_a{}^b \begin{pmatrix} \chi_1{}^{\dot{1}} & \chi_1{}^{\dot{2}} \\ \chi_2{}^{\dot{1}} & \chi_2{}^{\dot{2}} \end{pmatrix}_b{}^{\dot{e}} \begin{pmatrix} c & i\,s \\ i\,s & c \end{pmatrix}_{\dot{e}}{}^{T\ \dot{c}}$$

$$= \begin{pmatrix} c^2\chi_1{}^{\dot{1}} - s^2\chi_2{}^{\dot{2}} + i\,cs(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) & c^2\chi_1{}^{\dot{2}} - s^2\chi_2{}^{\dot{1}} + i\,cs(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}) \\ c^2\chi_2{}^{\dot{1}} - s^2\chi_1{}^{\dot{2}} + i\,cs(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}) & c^2\chi_2{}^{\dot{2}} - s^2\chi_1{}^{\dot{1}} + i\,cs(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) \end{pmatrix}$$

$$c \equiv \cos\left(\frac{\theta}{2}\right) \ , \quad s \equiv \sin\left(\frac{\theta}{2}\right)$$

Some trigonometric identities will be useful. Since $\cos(2x) = 2\cos^2 x - 1 = 1 - 2\sin^2 x$ and $\sin(2x) = 2\sin x \cos x$, we have $c^2 = \frac{1}{2}(1 + \cos\theta)$, $s^2 = \frac{1}{2}(1 - \cos\theta)$ and $cs = \frac{1}{2}\sin\theta$. With these, the above transformations simplify to:

$$\chi_1{}^{\dot{1}} \to \frac{1}{2}(1 + \cos\theta)\chi_1{}^{\dot{1}} - \frac{1}{2}(1 - \cos\theta)\chi_2{}^{\dot{2}} + \frac{i}{2}\sin\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}})$$

$$= \frac{1}{2}(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}}) + \frac{1}{2}\cos\theta(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}) + \frac{i}{2}\sin\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}})$$

$$\chi_1{}^{\dot{2}} \to \frac{1}{2}(1 + \cos\theta)\chi_1{}^{\dot{2}} - \frac{1}{2}(1 - \cos\theta)\chi_2{}^{\dot{1}} + \frac{i}{2}\sin\theta(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}})$$

$$= \frac{1}{2}(\chi_1{}^{\dot{2}} - \chi_2{}^{\dot{1}}) + \frac{1}{2}\cos\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) + \frac{i}{2}\sin\theta(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}})$$

$$\chi_2{}^{\dot{1}} \to \frac{1}{2}(1 + \cos\theta)\chi_2{}^{\dot{1}} - \frac{1}{2}(1 - \cos\theta)\chi_1{}^{\dot{2}} + \frac{i}{2}\sin\theta(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}})$$

$$= -\frac{1}{2}(\chi_1{}^{\dot{2}} - \chi_2{}^{\dot{1}}) + \frac{1}{2}\cos\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) + \frac{i}{2}\sin\theta(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}})$$

$$\chi_2{}^{\dot{2}} \to \frac{1}{2}(1 + \cos\theta)\chi_2{}^{\dot{2}} - \frac{1}{2}(1 - \cos\theta)\chi_1{}^{\dot{1}} + \frac{i}{2}\sin\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}})$$

$$= -\frac{1}{2}(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}}) + \frac{1}{2}\cos\theta(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}) + \frac{i}{2}\sin\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}})$$

These all appear in $\pm$ pairs, so add and subtract some of these transformations:

$$\chi_1{}^{\dot{2}} - \chi_2{}^{\dot{1}} \to \chi_1{}^{\dot{2}} - \chi_2{}^{\dot{1}}$$
$$\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}} \to \chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}}$$
$$\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}} \to \cos\theta(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}) + i\,\sin\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}})$$
$$\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}} \to \cos\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) + i\,\sin\theta(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}})$$

Define the symbols $(v^t, \vec{v})$ by $v^t \equiv \chi_2{}^{\dot{1}} - \chi_1{}^{\dot{2}}, v^x \equiv \chi_2{}^{\dot{2}} - \chi_1{}^{\dot{1}}, v^y \equiv -i(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}), v^z \equiv \chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}$.

43

In terms of these new labels, then above transformation laws are:

$$
\begin{pmatrix} v^t \\ v^x \\ v^y \\ v^z \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} v^t \\ v^x \\ v^y \\ v^z \end{pmatrix}
$$

That is exactly how a 4-vector v^μ transforms under rotations about the x-axis.

For $\hat{n} = (0,1,0)$, the transformation matrix is $e^{i\theta\,\hat{n}\cdot(\vec{\sigma}/2)} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$, so the components of χ transform as:

$$
\begin{aligned}
\chi_1{}^{\dot{1}} &\rightarrow c^2\chi_1{}^{\dot{1}} + s^2\chi_2{}^{\dot{2}} + cs(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) \\
&= \frac{1}{2}(1+\cos\theta)\chi_1{}^{\dot{1}} + \frac{1}{2}(1-\cos\theta)\chi_2{}^{\dot{2}} + \frac{1}{2}\sin\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) \\
&= \frac{1}{2}(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}) + \frac{1}{2}\cos\theta(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}}) + \frac{1}{2}\sin\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}})
\end{aligned}
$$

$$
\begin{aligned}
\chi_1{}^{\dot{2}} &\rightarrow c^2\chi_1{}^{\dot{2}} - s^2\chi_2{}^{\dot{1}} - cs(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}}) \\
&= \frac{1}{2}(1+\cos\theta)\chi_1{}^{\dot{2}} - \frac{1}{2}(1-\cos\theta)\chi_2{}^{\dot{1}} - \frac{1}{2}\sin\theta(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}}) \\
&= \frac{1}{2}(\chi_1{}^{\dot{2}} - \chi_2{}^{\dot{1}}) + \frac{1}{2}\cos\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) - \frac{1}{2}\sin\theta(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}})
\end{aligned}
$$

$$
\begin{aligned}
\chi_2{}^{\dot{1}} &\rightarrow c^2\chi_2{}^{\dot{1}} - s^2\chi_1{}^{\dot{2}} - cs(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}}) \\
&= \frac{1}{2}(1+\cos\theta)\chi_2{}^{\dot{1}} - \frac{1}{2}(1-\cos\theta)\chi_1{}^{\dot{2}} - \frac{1}{2}\sin\theta(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}}) \\
&= -\frac{1}{2}(\chi_1{}^{\dot{2}} - \chi_2{}^{\dot{1}}) + \frac{1}{2}\cos\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) - \frac{1}{2}\sin\theta(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}})
\end{aligned}
$$

$$
\begin{aligned}
\chi_2{}^{\dot{2}} &\rightarrow c^2\chi_2{}^{\dot{2}} + s^2\chi_1{}^{\dot{1}} - cs(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) \\
&= \frac{1}{2}(1+\cos\theta)\chi_2{}^{\dot{2}} + \frac{1}{2}(1-\cos\theta)\chi_1{}^{\dot{1}} - \frac{1}{2}\sin\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) \\
&= \frac{1}{2}(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}) - \frac{1}{2}\cos\theta(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}}) - \frac{1}{2}\sin\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}})
\end{aligned}
$$

Adding and subtracting these gives:

$$
\begin{aligned}
\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}} &\rightarrow \chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}} \\
\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}} &\rightarrow \cos\theta(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}}) + \sin\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) \\
\chi_1{}^{\dot{2}} - \chi_2{}^{\dot{1}} &\rightarrow \chi_1{}^{\dot{2}} - \chi_2{}^{\dot{1}} \\
\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}} &\rightarrow \cos\theta(\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}) - \sin\theta(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}})
\end{aligned}
$$

Using the same definitions from before, $v^t \equiv \chi_2{}^{\dot{1}} - \chi_1{}^{\dot{2}}, v^x \equiv \chi_2{}^{\dot{2}} - \chi_1{}^{\dot{1}}, v^y \equiv -i(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}), v^z \equiv \chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}$, we get:

$$\begin{pmatrix} v^t \\ v^x \\ v^y \\ v^z \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & -\sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & \sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} v^t \\ v^x \\ v^y \\ v^z \end{pmatrix}$$

That is exactly how a 4-vector v^μ transforms under rotations about the y-axis.

For $\hat{n} = (0, 0, 1)$, the transformation matrix is $e^{i\theta\,\hat{n}\cdot(\vec{\sigma}/2)} = \begin{pmatrix} c + is & 0 \\ 0 & c - is \end{pmatrix} = \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix}$,
so the transformation of χ is easy:

$$\chi_1{}^{\dot{1}} \rightarrow (\cos\theta + i\,\sin\theta)\chi_1{}^{\dot{1}}$$
$$\chi_1{}^{\dot{2}} \rightarrow \chi_1{}^{\dot{2}}$$
$$\chi_2{}^{\dot{1}} \rightarrow \chi_2{}^{\dot{1}}$$
$$\chi_2{}^{\dot{2}} \rightarrow (\cos\theta - i\,\sin\theta)\chi_2{}^{\dot{2}}$$

Rearrange these to make them look like the previous two cases:

$$\chi_1{}^{\dot{2}} - \chi_2{}^{\dot{1}} \rightarrow \chi_1{}^{\dot{2}} - \chi_2{}^{\dot{1}}$$
$$\chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}} \rightarrow \chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}$$
$$\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}} \rightarrow \cos\theta(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}) + i\,\sin\theta(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}})$$
$$\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}} \rightarrow \cos\theta(\chi_1{}^{\dot{1}} - \chi_2{}^{\dot{2}}) + i\,\sin\theta(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}})$$

So, again, with the same definitions $v^t \equiv \chi_2{}^{\dot{1}} - \chi_1{}^{\dot{2}}, v^x \equiv \chi_2{}^{\dot{2}} - \chi_1{}^{\dot{1}}, v^y \equiv -i(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}), v^z \equiv \chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}$, these transformations are:

$$\begin{pmatrix} v^t \\ v^x \\ v^y \\ v^z \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v^t \\ v^x \\ v^y \\ v^z \end{pmatrix}$$

That is exactly how a 4-vector v^μ transforms under rotations about the z-axis. Therefore, we know that the object $\chi_a{}^{\dot{c}}$, which transforms under the $(\frac{1}{2}, \frac{1}{2})$-representation of the Lorentz group, can be repackaged into an object v^μ, which transforms under the 4-dimensional ("vector") representation of the Lorentz group.

Before proceeding to check the boosts, let us reflect on the repackaging of the components of

χ. The definitions we used are:

$$v^t \equiv \chi_2{}^{\dot{1}} - \chi_1{}^{\dot{2}}$$
$$v^x \equiv \chi_2{}^{\dot{2}} - \chi_1{}^{\dot{1}}$$
$$v^y \equiv -i(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}})$$
$$v^z \equiv \chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}}$$

Note the following very important point: since $SU(2)$ indices are raised and lowered with the antisymmetric tensors $\epsilon^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\epsilon_{ab} = \begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix}$, the diagonal components of $\chi_a{}^{\dot{c}}$ correspond to the off-diagonal components of $\chi^{a\dot{c}}$ and vice versa, and all sorts of minus signs appear. For example, $\chi_2{}^{\dot{1}} = \epsilon_{2a}\chi^{a\dot{1}} = \epsilon_{21}\chi^{1\dot{1}} = -\epsilon_{12}\chi^{1\dot{1}} = +\epsilon^{12}\chi^{1\dot{1}} = \chi^{1\dot{1}}$. Raising all of the indices in that way results in the following for the definitions of $(v^t, \vec{v})$:

$$v^t \equiv \chi_2{}^{\dot{1}} - \chi_1{}^{\dot{2}} = +\chi^{1\dot{1}} + \chi^{2\dot{2}} = I_{a\dot{a}}\,\chi^{a\dot{a}}$$
$$v^x \equiv \chi_2{}^{\dot{2}} - \chi_1{}^{\dot{1}} = +\chi^{1\dot{2}} + \chi^{2\dot{1}} = \sigma^x_{a\dot{a}}\,\chi^{a\dot{a}}$$
$$v^y \equiv -i(\chi_1{}^{\dot{1}} + \chi_2{}^{\dot{2}}) = +i\chi^{2\dot{1}} - i\chi^{1\dot{2}} = \sigma^y_{a\dot{a}}\,\chi^{a\dot{a}}$$
$$v^z \equiv \chi_1{}^{\dot{2}} + \chi_2{}^{\dot{1}} = -\chi^{2\dot{2}} + \chi^{1\dot{1}} = \sigma^z_{a\dot{a}}\,\chi^{a\dot{a}}$$

We therefore have discovered a way to turn the pair $a\dot{a} \in \{1\dot{1}, 1\dot{2}, 2\dot{1}, 2\dot{2}\}$ of $SU(2)_L \otimes SU(2)_R$ indices into the single $SO(3,1)$ 4-vector index, $\mu \in \{0, 1, 2, 3\}$: for any value of $\mu \in \{x, y, z\}$, use the corresponding Pauli matrix to contract the $a\dot{a}$ indices; for $\mu = 0$, use the identity matrix. This is often written as $\sigma^\mu_{a\dot{a}}$, where $\sigma^0 \equiv I$.

Back to the problem at hand, we still have to check that the boosts work, but now we have two options as to how to go about doing that. We could proceed as before, by writing out the boost matrices explicitly, calculating the transformation of each component of χ, repackaging them into a 4-vector, and showing that it's the same 4-vector that we got before.

However, we now have a second option: if it is true that we can use the machine $\sigma^\mu_{a\dot{a}}$ to change the $(\frac{1}{2}, \frac{1}{2})$ indices to one 4-index, then it is already manifest that the object $\chi^\mu \equiv \sigma^\mu_{a\dot{a}}\chi^{a\dot{a}}$ transforms as a 4-vector under rotations and boosts, and therefore that the two representations are equivalent. A valid way to approach this problem is simply to check that the symbol $\sigma^\mu_{a\dot{a}}$ is invariant under simultaneous Lorentz transformations for all of its indices.

It's easiest to work with an object that has one lower $SU(2)_L$ index and one upper $SU(2)_R$ index, so define the quantity $s_a^{\mu\ \dot{c}} \equiv -\sigma^\mu_{a\dot{a}}\epsilon^{\dot{a}\dot{c}}$ whose components are, numerically:

$$s_a^{t\ \dot{c}} = (-i\sigma^y)_a{}^{\dot{c}}$$
$$s_a^{x\ \dot{c}} = (\sigma^z)_a{}^{\dot{c}}$$
$$s_a^{y\ \dot{c}} = (-i\sigma^t)_a{}^{\dot{c}}$$
$$s_a^{z\ \dot{c}} = (-\sigma^x)_a{}^{\dot{c}}$$

46

Let's take a boost in the x-direction, meaning $\hat{n} = (1, 0, 0)$, and apply it to $s_a^{\mu\ \dot{c}}$:

$$s_a^{\mu\ \dot{c}} \rightarrow \left[e^{\frac{\beta}{2}\sigma^x} \right]_a^{\ b} \left[e^{-\frac{\beta}{2}\sigma^x} \right]_{\dot{e}}^{\ \dot{c}} \Lambda^\mu_{\ \nu}\, s_b^{\nu\ \dot{e}}$$

$$= \begin{pmatrix} \cosh\frac{\beta}{2} & \sinh\frac{\beta}{2} \\ \sinh\frac{\beta}{2} & \cosh\frac{\beta}{2} \end{pmatrix}_a^{\ b} \begin{pmatrix} \cosh\frac{\beta}{2} & -\sinh\frac{\beta}{2} \\ -\sinh\frac{\beta}{2} & \cosh\frac{\beta}{2} \end{pmatrix}_{\dot{e}}^{\ \dot{c}} \begin{pmatrix} \cosh\beta & -\sinh\beta & 0 & 0 \\ -\sinh\beta & \cosh\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_\nu^{\mu}\, s_b^{\nu\ \dot{e}}$$

Do one index at a time. For $\mu = t$, we have:

$$\Lambda^t_{\ \nu}\, s_b^{\nu\ \dot{e}} = \cosh\beta (i\sigma^y)_b^{\ \dot{e}} - \sinh\beta (\sigma^z)_b^{\ \dot{e}}$$

$$= \cosh\beta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}_b^{\ \dot{e}} - \sinh\beta \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_b^{\ \dot{e}}$$

$$= \begin{pmatrix} -\sinh\beta & -\cosh\beta \\ \cosh\beta & \sinh\beta \end{pmatrix}_b^{\ \dot{e}}$$

Therefore,

$$s_a^{t\ \dot{c}} \rightarrow \left[\begin{pmatrix} \cosh\frac{\beta}{2} & \sinh\frac{\beta}{2} \\ \sinh\frac{\beta}{2} & \cosh\frac{\beta}{2} \end{pmatrix} \begin{pmatrix} -\sinh\beta & -\cosh\beta \\ \cosh\beta & \sinh\beta \end{pmatrix} \begin{pmatrix} \cosh\frac{\beta}{2} & -\sinh\frac{\beta}{2} \\ -\sinh\frac{\beta}{2} & \cosh\frac{\beta}{2} \end{pmatrix} \right]_a^{\ \dot{c}}$$

$$= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}_a^{\ \dot{c}} = (-i\sigma^y)_a^{\ \dot{c}} = s_a^{t\ \dot{c}}\quad \checkmark$$

For $\mu = x$, we have:

$$\Lambda^x_{\ \nu} s_b^{\nu\ \dot{e}} = -\sinh\beta (-i\sigma^y)_b^{\ \dot{e}} + \cosh\beta (\sigma^z)_b^{\ \dot{e}}$$

$$= -\sinh\beta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}_b^{\ \dot{e}} + \cosh\beta \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_b^{\ \dot{e}}$$

$$= \begin{pmatrix} \cosh\beta & \sinh\beta \\ -\sinh\beta & -\cosh\beta \end{pmatrix}_b^{\ \dot{e}}$$

Therefore,

$$s_a^{x\ \dot{c}} \rightarrow \left[\begin{pmatrix} \cosh\frac{\beta}{2} & \sinh\frac{\beta}{2} \\ \sinh\frac{\beta}{2} & \cosh\frac{\beta}{2} \end{pmatrix} \begin{pmatrix} \cosh\beta & \sinh\beta \\ -\sinh\beta & -\cosh\beta \end{pmatrix} \begin{pmatrix} \cosh\frac{\beta}{2} & -\sinh\frac{\beta}{2} \\ -\sinh\frac{\beta}{2} & \cosh\frac{\beta}{2} \end{pmatrix} \right]_a^{\ \dot{c}}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}_a^{\ \dot{c}} = (\sigma^z)_a^{\ \dot{c}} = s_a^{x\ \dot{c}}\quad \checkmark$$

For $\mu = y$, we have:

$$\Lambda^y_{\ \nu} s_b^{\nu\ \dot{e}} = (-i\sigma^t)_b^{\ \dot{e}} = \begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix}_b^{\ \dot{e}}$$

47

Therefore,

$$
s_a^{y\ \dot{c}} \rightarrow \left[\begin{pmatrix} \cosh\frac{\beta}{2} & \sinh\frac{\beta}{2} \\ \sinh\frac{\beta}{2} & \cosh\frac{\beta}{2} \end{pmatrix} \begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} \cosh\frac{\beta}{2} & -\sinh\frac{\beta}{2} \\ -\sinh\frac{\beta}{2} & \cosh\frac{\beta}{2} \end{pmatrix} \right]_a^{\dot{c}}
$$

$$
= \begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix}_a^{\dot{c}} = (-i\sigma^t)_a^{\dot{c}} = s_a^{y\ \dot{c}} \quad \checkmark
$$

For $\mu = z$, we have:

$$
\Lambda^z_{\ \nu} s_b^{\nu\ \dot{e}} = (-\sigma^x)_b^{\dot{e}} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}_b^{\dot{e}}
$$

Therefore,

$$
s_a^{z\ \dot{c}} \rightarrow \left[\begin{pmatrix} \cosh\frac{\beta}{2} & \sinh\frac{\beta}{2} \\ \sinh\frac{\beta}{2} & \cosh\frac{\beta}{2} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \cosh\frac{\beta}{2} & -\sinh\frac{\beta}{2} \\ -\sinh\frac{\beta}{2} & \cosh\frac{\beta}{2} \end{pmatrix} \right]_a^{\dot{c}}
$$

$$
= \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}_a^{\dot{c}} = (-\sigma^x)_a^{\dot{c}} = s_a^{z\ \dot{c}} \quad \checkmark
$$

Since ϵ^{ab} is invariant under $SU(2)_L$ (if it weren't, we couldn't use it to raise and lower $SU(2)_L$ indices), the invariance of $s_a^{\mu\ \dot{c}}$ implies the invariance of $\sigma^\mu_{a\dot{c}}$. We have checked only the case of a Lorentz boost in the x-direction, but since we have already checked for rotational invariance, we can just argue by symmetry that boosts in the y- and z- directions also leave the symbol $\sigma^\mu_{a\dot{c}}$ invariant. Therefore, the $(\frac{1}{2}, \frac{1}{2})$-representation of $SU(2)_L \otimes SU(2)_R$ is equivalent to the 4-vector representation of $SO(3, 1)$.

2. Work out how the six objects contained in the $(1, 0)$ and $(0, 1)$ transform under the Lorentz group. Recall from your course on electromagnetism how the electric and magnetic fields $\vec{E}$ and $\vec{B}$ transform. Conclude that the electromagnetic field in fact transforms as $(1, 0) \oplus (0, 1)$. Show that it is parity that once again forces us to use a reducible representation.

Solution:

A spin-j object transforms under the $(2j + 1)$-dimensional representation of $SU(2)$, so a spin-1 object transforms under the 3-dimensional representation of $SU(2)$. Let $a \in \{1, 2\}$ be an index for the 2-dimensional representation of $SU(2)$. An object X transforming under the $2 \otimes 2$-dimensional representation of $SU(2)$ therefore carries two indices: $X = X_{ab}$. The 3-dimensional representation of $SU(2)$ can be obtained from the symmetrized tensor product of two 2-dimensional representations, or in fancy notation, $3 = 2 \otimes_S 2$. Therefore, the quantity $X_{(ab)} \equiv \frac{1}{2}(X_{ab} + X_{ba})$ transforms under the 3-dimensional representation of $SU(2)$.

Now we can find out how the object $X_{(ab)}$ in the 3-dimensional representation of $SU(2)$ transforms, since each index gets its own transformation matrix. Explicitly, for a rotation through an angle θ about an axis $\hat{n}$, the transformation is:

$$X_{(ab)} \to \left[e^{i\theta\,\hat{n}\cdot\left(\frac{\vec{\sigma}}{2}\right)}\right]_a^{\ c} \left[e^{i\theta\,\hat{n}\cdot\left(\frac{\vec{\sigma}}{2}\right)}\right]_b^{\ d} X_{(cd)}$$

$$= \left[e^{i\theta\,\hat{n}\cdot\left(\frac{\vec{\sigma}}{2}\right)}\right]_a^{\ c} X_{(cd)} \left[e^{i\theta\,\hat{n}\cdot\left(\frac{\vec{\sigma}}{2}\right)}\right]_b^{T\,d}$$

$$= \begin{pmatrix} c + in^z s & in^- s \\ in^+ s & c - in^z s \end{pmatrix} \begin{pmatrix} X_{11} & X_{(12)} \\ X_{(12)} & X_{22} \end{pmatrix} \begin{pmatrix} c + in^z s & in^+ s \\ in^- s & c - in^z s \end{pmatrix}$$

$$\text{where} \quad c \equiv \cos\left(\frac{\theta}{2}\right) \ , \ s \equiv \sin\left(\frac{\theta}{2}\right)$$

For a boost through β about $\hat{n}$, the transformation is:

$$X_{(ab)} \to \begin{pmatrix} c_h + n^z s_h & n^- s_h \\ n^+ s_h & c_h - n^z s_h \end{pmatrix} \begin{pmatrix} X_{11} & X_{(12)} \\ X_{(12)} & X_{22} \end{pmatrix} \begin{pmatrix} c_h + n^z s_h & n^+ s_h \\ n^- s_h & c_h - n^z s_h \end{pmatrix}$$

$$c_h \equiv \cosh\left(\frac{\beta}{2}\right) \ , \ s_h \equiv \sinh\left(\frac{\beta}{2}\right)$$

We've just been referencing $SU(2)$, but remember that we're talking about the Lorentz group, $SO(3,1) \cong SU(2)_L \otimes SU(2)_R$. The discussion above applies for $SU(2)_L$, so the transformations we found are those for an object in the $(1,0)$-representation, meaning that said object transforms under the 3-dimensional representation of $SU(2)_L$ and is unaffected by $SU(2)_R$. What about the $(0,1)$-representation?

As for the case of $(0,\frac{1}{2})$ versus $(\frac{1}{2},0)$ from the previous problem, the $(0,1)$ transforms in exactly the same way as the $(1,0)$ except for the critical difference that $\vec{K}_{(1,0)} = +\vec{\sigma}/(2i)$ whereas $\vec{K}_{(0,1)} = -\vec{\sigma}/(2i)$. In practice that means we can copy the results from above except with $\beta \to -\beta$, which implies $c_h \to c_h$ and $s_h \to -s_h$.

Let Y transform under the $(0,1)$ representation of the Lorentz group: $Y = Y^{(\dot{a}\dot{b})}$. The Lorentz group transformations for Y are:

Rotation through θ about $\hat{n}$:

$$Y^{(\dot{a}\dot{b})} \to \begin{pmatrix} c + in^z s & in^- s \\ in^+ s & c - in^z s \end{pmatrix} \begin{pmatrix} Y^{\dot{1}\dot{1}} & Y^{(\dot{1}\dot{2})} \\ Y^{(\dot{1}\dot{2})} & Y^{\dot{2}\dot{2}} \end{pmatrix} \begin{pmatrix} c + in^z s & in^+ s \\ in^- s & c - in^z s \end{pmatrix}$$

$$c \equiv \cos\left(\frac{\theta}{2}\right) \ , \ s \equiv \sin\left(\frac{\theta}{2}\right)$$

Boost through β about $\hat{n}$:

$$Y^{(\dot{a}\dot{b})} \to \begin{pmatrix} c_h - n^z s_h & -n^- s_h \\ -n^+ s_h & c_h + n^z s_h \end{pmatrix} \begin{pmatrix} Y^{\dot{1}\dot{1}} & Y^{(\dot{1}\dot{2})} \\ Y^{(\dot{1}\dot{2})} & Y^{\dot{2}\dot{2}} \end{pmatrix} \begin{pmatrix} c_h - n^z s_h & -n^+ s_h \\ -n^- s_h & c_h + n^z s_h \end{pmatrix}$$

$$c_h \equiv \cosh\left(\frac{\beta}{2}\right) \ , \ s_h \equiv \sinh\left(\frac{\beta}{2}\right)$$

Now consider a quantity Φ that transforms under the $(1,0) \oplus (0,1)$ representation of the Lorentz group defined by the following:

$$\Phi \equiv X \oplus Y = \begin{pmatrix} X_{(ab)} & 0_{2\times 2} \\ 0_{2\times 2} & Y^{(\dot a \dot b)} \end{pmatrix} \equiv \begin{pmatrix} \Phi_1 & \Phi_3 & 0 & 0 \\ \Phi_3 & \Phi_2 & 0 & 0 \\ 0 & 0 & \Phi_4 & \Phi_6 \\ 0 & 0 & \Phi_6 & \Phi_5 \end{pmatrix}$$

This quantity Φ is a 4×4 matrix with 6 independent components. Using the above work, it transforms in the following way under boosts:

$\Phi \to$

$$\begin{pmatrix} c_h + n^z s_h & n^- s_h & 0 & 0 \\ n^+ s_h & c_h - n^z s_h & 0 & 0 \\ 0 & 0 & c_h - n^z s_h & -n^- s_h \\ 0 & 0 & -n^+ s_h & c_h + n^z s_h \end{pmatrix}\begin{pmatrix} \Phi_1 & \Phi_3 & 0 & 0 \\ \Phi_3 & \Phi_2 & 0 & 0 \\ 0 & 0 & \Phi_4 & \Phi_6 \\ 0 & 0 & \Phi_6 & \Phi_5 \end{pmatrix}\begin{pmatrix} c_h + n^z s_h & n^+ s_h & 0 & 0 \\ n^- s_h & c_h - n^z s_h & 0 & 0 \\ 0 & 0 & c_h - n^z s_h & -n^+ s_h \\ 0 & 0 & -n^- s_h & c_h + n^z s_h \end{pmatrix}$$

To understand what's going on, specialize to a boost in the x-direction: $\hat n = (1,0,0)$. For that case, the field Φ transforms as:

$$\Phi \to \begin{pmatrix} c_h & s_h & 0 & 0 \\ s_h & c_h & 0 & 0 \\ 0 & 0 & c_h & -s_h \\ 0 & 0 & -s_h & c_h \end{pmatrix}\begin{pmatrix} \Phi_1 & \Phi_3 & 0 & 0 \\ \Phi_3 & \Phi_2 & 0 & 0 \\ 0 & 0 & \Phi_4 & \Phi_6 \\ 0 & 0 & \Phi_6 & \Phi_5 \end{pmatrix}\begin{pmatrix} c_h & s_h & 0 & 0 \\ s_h & c_h & 0 & 0 \\ 0 & 0 & c_h & -s_h \\ 0 & 0 & -s_h & c_h \end{pmatrix}$$

Therefore, each component of Φ transforms as:

$$\Phi_1 \to c_h(c_h\Phi_1 + s_h\Phi_3) + s_h(c_h\Phi_3 + s_h\Phi_2)$$
$$\Phi_2 \to c_h(c_h\Phi_2 + s_h\Phi_3) + s_h(c_h\Phi_3 + s_h\Phi_1)$$
$$\Phi_3 \to c_h(c_h\Phi_3 + s_h\Phi_2) + s_h(c_h\Phi_1 + s_h\Phi_3)$$
$$\Phi_4 \to c_h(c_h\Phi_4 - s_h\Phi_6) - s_h(c_h\Phi_6 - s_h\Phi_5)$$
$$\Phi_5 \to c_h(c_h\Phi_5 - s_h\Phi_6) - s_h(c_h\Phi_6 - s_h\Phi_4)$$
$$\Phi_6 \to c_h(c_h\Phi_6 - s_h\Phi_5) - s_h(c_h\Phi_4 - s_h\Phi_6)$$

Now use some hyperbolic trigonometric identities. Recall that $c_h \equiv \cosh(\beta/2)$ and $s_h \equiv \sinh(\beta/2)$, so that $c_h^2 = \frac{1}{2}(\cosh\beta + 1)$, $s_h^2 = \frac{1}{2}(\cosh\beta - 1)$, and $2c_h s_h = \sinh\beta$. Using these, we get:

$$\Phi_1 \to c_h^2\Phi_1 + s_h^2\Phi_2 + 2s_h c_h\Phi_3$$
$$= \frac{1}{2}(\cosh\beta + 1)\Phi_1 + \frac{1}{2}(\cosh\beta - 1)\Phi_2 + \sinh\beta \ \Phi_3$$
$$= \cosh\beta \ \frac{1}{2}(\Phi_1 + \Phi_2) + \frac{1}{2}(\Phi_1 - \Phi_2) + \sinh\beta \ \Phi_3$$

50

$$\Phi_2 \rightarrow c_h^2 \Phi_2 + s_h^2 \Phi_1 + 2s_h c_h \Phi_3$$
$$= \frac{1}{2}(\cosh\beta + 1)\Phi_2 + \frac{1}{2}(\cosh\beta - 1)\Phi_1 + \sinh\beta\ \Phi_3$$
$$= \cosh\beta\ \frac{1}{2}(\Phi_1 + \Phi_2) - \frac{1}{2}(\Phi_1 - \Phi_2) + \sinh\beta\ \Phi_3$$

$$\Phi_3 \rightarrow (c_h^2 + s_h^2)\Phi_3 + c_h s_h(\Phi_1 + \Phi_2)$$
$$= \cosh\beta\ \Phi_3 + \sinh\beta\ \frac{1}{2}(\Phi_1 + \Phi_2)$$

The fields appear in the combinations $\Phi_\pm \equiv \frac{1}{2}(\Phi_1 \pm \Phi_2)$ and Φ_3. From adding and subtracting the transformations for Φ_1 and Φ_2, we get:

$$\Phi_+ \rightarrow \cosh\beta\ \Phi_+ + \sinh\beta\ \Phi_3$$
$$\Phi_3 \rightarrow \cosh\beta\ \Phi_3 + \sinh\beta\ \Phi_+$$
$$\Phi_- \rightarrow \Phi_-$$

Now do the same thing for Φ_4, Φ_5 and Φ_6:

$$\Phi_4 \rightarrow c_h^2 \Phi_4 + s_h^2 \Phi_5 - 2c_h s_h \Phi_6$$
$$= \cosh\beta\ \frac{1}{2}(\Phi_4 + \Phi_5) + \frac{1}{2}(\Phi_4 - \Phi_5) - \sinh\beta\ \Phi_6$$

$$\Phi_5 \rightarrow c_h^2 \Phi_5 + s_h^2 \Phi_4 - 2c_h s_h \Phi_6$$
$$= \cosh\beta\ \frac{1}{2}(\Phi_4 + \Phi_5) - \frac{1}{2}(\Phi_4 - \Phi_5) - \sinh\beta\ \Phi_6$$

$$\Phi_6 \rightarrow (c_h^2 + s_h^2)\Phi_6 - c_h s_h(\Phi_4 + \Phi_5)$$
$$= \cosh\beta\ \Phi_6 - \sinh\beta\ \frac{1}{2}(\Phi_4 + \Phi_5)$$

As before, define $\tilde{\Phi}_\pm \equiv \frac{1}{2}(\Phi_4 \pm \Phi_5)$ so that the transformations are:

$$\tilde{\Phi}_+ \rightarrow \cosh\beta\ \tilde{\Phi}_+ - \sinh\beta\ \Phi_6$$
$$\Phi_6 \rightarrow \cosh\beta\ \Phi_6 - \sinh\beta\ \tilde{\Phi}_+$$
$$\tilde{\Phi}_- \rightarrow \tilde{\Phi}_-$$

To collect and rearrange the results, boosting in the x-direction with speed $v = \tanh\beta$

transforms the components of Φ as follows:

$$\Phi_- \to \Phi_-$$
$$\Phi_+ \to \cosh\beta\ \Phi_+ + \sinh\beta\ \Phi_3$$
$$\Phi_6 \to \cosh\beta\ \Phi_6 - \sinh\beta\ \tilde{\Phi}_+$$
$$\tilde{\Phi}_- \to \tilde{\Phi}_-$$
$$\tilde{\Phi}_+ \to \cosh\beta\ \tilde{\Phi}_+ - \sinh\beta\ \Phi_6$$
$$\Phi_3 \to \cosh\beta\ \Phi_3 + \sinh\beta\ \Phi_+$$

With the relabeling $E^x \equiv \Phi_-, E^y \equiv \Phi_+, E^z \equiv \Phi_6, B^x \equiv \tilde{\Phi}_-, B^y \equiv \tilde{\Phi}_+$ and $B^z \equiv \Phi_3$, we get:

$$E^x \to E^x$$
$$E^y \to \cosh\beta\ E^y + \sinh\beta\ B^z$$
$$E^z \to \cosh\beta\ E^z - \sinh\beta\ B^y$$
$$B^x \to B^x$$
$$B^y \to \cosh\beta\ B^y - \sinh\beta\ E^z$$
$$B^z \to \cosh\beta\ B^z + \sinh\beta\ E^y$$

These are precisely the transformation rules of an electric field $\vec{E}$ and a magnetic field $\vec{B}$ under a Lorentz boost in the x-direction with speed $v = \tanh\beta$. We see that the electromagnetic field transforms under the reducible representation $(1,0) \oplus (0,1)$ of the Lorentz group.

That fact is typically repackaged in a different way. A spin-1 representation corresponds to the 3-dimensional representation of the rotation group, which corresponds to the 4-dimensional representation of the Lorentz group, denoted by the index $\mu\ \epsilon\ \{0,1,2,3\}$. Therefore, an object that transforms under two copies of spin-1 implies that the object gets two such indices, μ and ν.

But a matrix $M^{\mu\nu}$ has $4 \times 4 = 16$ independent components, where as the object Φ has only 6 independent components. However, the antisymmetric matrix $F^{\mu\nu} \equiv \frac{1}{2}(M^{\mu\nu} - M^{\nu\mu})$ has $4 \times (4-1)/2 = 6$ independent components. Such a matrix F transforms under a Lorentz boost with speed $v = \tanh\beta$ in the x-direction as:

$$F^{\mu\nu} \to \begin{pmatrix} \cosh\beta & \sinh\beta & 0 & 0 \\ \sinh\beta & \cosh\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & F^{01} & F^{02} & F^{03} \\ -F^{01} & 0 & F^{12} & F^{13} \\ -F^{02} & -F^{12} & 0 & F^{23} \\ -F^{03} & -F^{13} & -F^{23} & 0 \end{pmatrix} \begin{pmatrix} \cosh\beta & \sinh\beta & 0 & 0 \\ \sinh\beta & \cosh\beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^T$$

So the 6 components transform as:

$$F^{01} \rightarrow F^{01}$$
$$F^{02} \rightarrow \cosh\beta \; F^{02} + \sinh\beta \; F^{12}$$
$$F^{03} \rightarrow \cosh\beta \; F^{03} + \sinh\beta \; F^{13}$$
$$F^{23} \rightarrow F^{23}$$
$$F^{13} \rightarrow \cosh\beta \; F^{13} + \sinh\beta \; F^{03}$$
$$F^{12} \rightarrow \cosh\beta \; F^{12} + \sinh\beta \; F^{02}$$

Comparing with the transformations for the components of $\vec{E}$ and $\vec{B}$ implies $F^{01} = E^x$, $F^{02} = E^y$, $F^{03} = E^z$, $F^{23} = B^x$, $F^{13} = -B^y$ and $F^{12} = B^z$, or more compactly, $E^i = F^{0i}$ and $B^i = \frac{1}{2}\epsilon^{ijk}F_{jk}$.

We know that parity interchanges $SU(2)_L$ with $SU(2)_R$. How can we get an object that transforms as spin-1 under the rotation group (that is, as a "3-vector" under spatial rotations) that is also invariant under parity? As discussed, an object that transforms as $(1,0)$ of $SU(2)_L \otimes SU(2)_R$ does in fact transform as spin-1 under the rotation group, but a parity operation would change such an object to $(0,1)$. If we want the theory of the electromagnetic field to preserve parity and to be Lorentz invariant, then it must transform as $(1,0) \oplus (0,1)$ under Lorentz transformations.

3. Show that
$$e^{-i\vec{\varphi}\cdot\vec{K}}\gamma^0 e^{+i\vec{\varphi}\cdot\vec{K}} = \begin{pmatrix} 0 & e^{-\vec{\varphi}\cdot\vec{\sigma}} \\ e^{+\vec{\varphi}\cdot\vec{\sigma}} & 0 \end{pmatrix}$$

and
$$e^{\vec{\varphi}\cdot\vec{\sigma}} = \cosh\varphi + \vec{\sigma}\cdot\hat{\varphi}\sinh\varphi$$

with the unit vector $\hat{\varphi} \equiv \vec{\varphi}/\varphi$. Identifying $\vec{p} = m\hat{\varphi}\sinh\varphi$, derive the Dirac equation. Show that
$$\gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}.$$

Solution:

If D is a diagonal matrix, then $e^D = \mathrm{diag}(e^{D_{11}}, e^{D_{22}}, ...)$. Page 117 tells us that

$$i\vec{K} = \begin{pmatrix} \frac{\vec{\sigma}}{2} & 0 \\ 0 & -\frac{\vec{\sigma}}{2} \end{pmatrix} \quad \text{and} \quad \gamma^0 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

so by direct computation we have

$$e^{-\vec{\varphi}\cdot(i\vec{K})}\gamma^0 e^{+\vec{\varphi}\cdot(i\vec{K})} = e^{-\vec{\varphi}\cdot\begin{pmatrix}\vec{\sigma}/2 & 0 \\ 0 & -\vec{\sigma}/2\end{pmatrix}}\gamma^0 e^{+\vec{\varphi}\cdot\begin{pmatrix}\vec{\sigma}/2 & 0 \\ 0 & -\vec{\sigma}/2\end{pmatrix}}$$

$$= \begin{pmatrix} e^{-\vec{\varphi}\cdot(\vec{\sigma}/2)} & 0 \\ 0 & e^{+\vec{\varphi}\cdot(\vec{\sigma}/2)} \end{pmatrix}\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}\begin{pmatrix} e^{+\vec{\varphi}\cdot(\vec{\sigma}/2)} & 0 \\ 0 & e^{-\vec{\varphi}\cdot(\vec{\sigma}/2)} \end{pmatrix}$$

$$= \begin{pmatrix} e^{-\vec{\varphi}\cdot(\vec{\sigma}/2)} & 0 \\ 0 & e^{+\vec{\varphi}\cdot(\vec{\sigma}/2)} \end{pmatrix}\begin{pmatrix} 0 & e^{-\vec{\varphi}\cdot(\vec{\sigma}/2)} \\ e^{+\vec{\varphi}\cdot(\vec{\sigma}/2)} & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & e^{-\vec{\varphi}\cdot\vec{\sigma}} \\ e^{+\vec{\varphi}\cdot\vec{\sigma}} & 0 \end{pmatrix}$$

Next, compute $e^{i(\theta\,\hat{\varphi}\cdot\vec{\sigma})} = \cos(\theta\,\hat{\varphi}\cdot\vec{\sigma}) + i\sin(\theta\,\hat{\varphi}\cdot\vec{\sigma})$ by Taylor expansion:

$$\cos(\theta\,\hat{\varphi}\cdot\vec{\sigma}) = \sum_{n=0,2,4,6...}^{\infty}\frac{(-1)^n}{n!}(\theta\,\hat{\varphi}\cdot\vec{\sigma})^n$$

Since $\{\sigma^i, \sigma^j\} = 2\delta^{ij}I$, any even power of $(\hat{\varphi}\cdot\vec{\sigma})$ is equal to the identity matrix:

$$(\hat{\varphi}\cdot\vec{\sigma})^2 = \hat{\varphi}_i\hat{\varphi}_j\sigma^i\sigma^j = \frac{1}{2}\hat{\varphi}_i\hat{\varphi}_j\{\sigma^i, \sigma^j\} = \hat{\varphi}_i\hat{\varphi}^i I = \hat{\varphi}\cdot\hat{\varphi}I = I$$

Therefore we have the simple result $\cos(\theta\,\hat{\varphi}\cdot\vec{\sigma}) = I\cos\theta$.

By the same argument, any odd power of $(\hat{\varphi}\cdot\vec{\sigma})$ is equal to itself. This allows us to write another simple expression: $\sin(\theta\,\hat{\varphi}\cdot\vec{\sigma}) = \hat{\varphi}\cdot\vec{\sigma}\sin\theta$. Therefore, we have just proved that

$$e^{i(\theta\,\hat{\varphi}\cdot\vec{\sigma})} = I\cos\theta + i\hat{\varphi}\cdot\vec{\sigma}\sin\theta$$

Since $\sin(ix) = i\sinh x$, we can let $\theta = -i\varphi$ to get the desired result

$$e^{\vec{\varphi}\cdot\vec{\sigma}} = I\cosh\varphi + \vec{\varphi}\cdot\vec{\sigma}\sinh\varphi$$

Next, we start with the form of the boosted Dirac equation written above equation (17) on p. 118:

$$\left(e^{-i\vec{\varphi}\cdot\vec{K}}\gamma^0 e^{+i\vec{\varphi}\cdot\vec{K}} - I_4\right)\psi(p) = 0$$

Using what we just derived in the first part of this problem gives

$$\begin{pmatrix} -I_2 & e^{-\vec{\varphi}\cdot\vec{\sigma}} \\ e^{+\vec{\varphi}\cdot\vec{\sigma}} & -I_2 \end{pmatrix}\psi(p) = 0$$

Using the second part of this problem, the above becomes

$$\begin{pmatrix} -I_2 & I_2\cosh\varphi - \vec{\varphi}\cdot\vec{\sigma}\sinh\varphi \\ I_2\cosh\varphi + \vec{\varphi}\cdot\vec{\sigma}\sinh\varphi & -I_2 \end{pmatrix}\psi(p) = 0$$

54

Identifying $\vec{p} = \hat{\varphi} m \sinh\varphi$ gives $\cosh\varphi = \sqrt{1 + |\vec{p}|^2/m^2} = \sqrt{m^2 + |\vec{p}|^2}/m = p^0/m$, so after multiplying through by m, the above equation becomes

$$\begin{pmatrix} -m & I_2 p^0 - \vec{\sigma}\cdot\vec{p} \\ I_2 p^0 + \vec{\sigma}\cdot\vec{p} & -m \end{pmatrix} \psi(p) = 0$$

Define the 4-vectors $\sigma^\mu \equiv (I_2, \vec{\sigma})$ and $\bar{\sigma}^\mu \equiv (I_2, -\vec{\sigma})$ (the bar is part of the name and does not denote any sort of complex conjugation operation) to rewrite the above as

$$\begin{pmatrix} -m & \sigma^\mu p_\mu \\ \bar{\sigma}^\mu p_\mu & -m \end{pmatrix} \psi(p) = 0$$

Defining $\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}$ gives the Dirac equation:

$$(\gamma^\mu p_\mu - m)\psi(p) = 0$$

This shows that $\gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}$, which completes the problem.

II.4 Spin-Statistics Connection

1. Show that we would also get into trouble if we quantize the Dirac field with commutation instead of anticommutation rules. Calculate the commutator $[J^0(\vec{x}, 0),\ J^0(\vec{0}, 0)]$.

Solution:

The charge density is $J^0 = \bar{\psi}\gamma^0\psi = \psi^\dagger\psi$. Consider the commutation relations

$$[\psi_\alpha(\vec{x}, t), \psi_\beta^\dagger(\vec{0}, t)]_\zeta = \delta^{(3)}(\vec{x})\delta_{\alpha\beta} \ \text{ and } \ [\psi_\alpha(\vec{x}, t), \psi_\beta(\vec{0}, t)]_\zeta = 0$$

where $[A, B]_\zeta \equiv AB - \zeta BA$ denotes the commutator for $\zeta = +1$ and the anticommutator for $\zeta = -1$. From now on all times t will be set to zero, so for notational convenience drop the vector arrow over the spatial coordinates. The commutator of charge densities is

$$
\begin{aligned}
[J^0(x), J^0(0)] &= [\psi^\dagger(x)\psi(x), \psi^\dagger(0)\psi(0)] \\
&= \psi_\alpha^\dagger(x)\psi_\alpha(x)\psi_\beta^\dagger(0)\psi_\beta(0) - \psi_\beta^\dagger(0)\psi_\beta(0)\psi_\alpha^\dagger(x)\psi_\alpha(x) \\
&= \psi_\alpha^\dagger(x)\psi_\alpha(x)\psi_\beta^\dagger(0)\psi_\beta(0) - \psi_\beta^\dagger(0)[\zeta\psi_\alpha^\dagger(x)\psi_\beta(0) + \delta(x)\delta_{\alpha\beta}]\psi_\alpha(x) \\
&= \psi_\alpha^\dagger(x)\psi_\alpha(x)\psi_\beta^\dagger(0)\psi_\beta(0) - \zeta\psi_\beta^\dagger(0)\psi_\alpha^\dagger(x)\psi_\beta(0)\psi_\alpha(x) - \zeta\delta(x)\psi_\alpha^\dagger(0)\psi_\alpha(x) \\
&= \psi_\alpha^\dagger(x)\psi_\alpha(x)\psi_\beta^\dagger(0)\psi_\beta(0) - \zeta[\zeta\psi_\alpha^\dagger(x)\psi_\beta^\dagger(0)][\zeta\psi_\alpha(x)\psi_\beta(0)] - \zeta\delta(x)\psi_\alpha^\dagger(0)\psi_\alpha(x) \\
&= \psi_\alpha^\dagger(x)\psi_\alpha(x)\psi_\beta^\dagger(0)\psi_\beta(0) - \zeta\psi_\alpha^\dagger(x)[\zeta\psi_\alpha(x)\psi_\beta^\dagger(0) + \delta(x)\delta_{\alpha\beta}]\psi_\beta(0) - \zeta\delta(x)\psi_\alpha^\dagger(0)\psi_\alpha(x) \\
&= \psi_\alpha^\dagger(x)\psi_\alpha(x)\psi_\beta^\dagger(0)\psi_\beta(0) - \psi_\alpha^\dagger(x)\psi_\alpha(x)\psi_\beta^\dagger(0)\psi_\beta(0) - \delta(x)[\psi_\alpha^\dagger(x)\psi_\alpha(0) + \zeta\psi_\alpha^\dagger(0)\psi_\alpha(x)] \\
&= -\delta(x)[\psi_\alpha^\dagger(x)\psi_\alpha(0) + \zeta\psi_\alpha^\dagger(0)\psi_\alpha(x)]
\end{aligned}
$$

The delta function is only nonzero when $x = 0$, so we have

$$[J^0(x), J^0(0)] = -\delta(x)\psi_\alpha^\dagger(0)\psi_\alpha(0)(1 + \zeta) .$$

If $\zeta = -1$, meaning fermion anticommutation relations, then $[J^0(x), J^0(0)] = 0$ as desired. But if $\zeta = +1$, meaning fermion commutation relations, then indeed we run into trouble.

II.5 Vacuum Energy, Grassmann Integrals, and Feynman Diagrams for Fermions

1. Write down the Feynman amplitude for the diagram in figure II.5.1 for the scalar theory (19). The answer is given in chapter III.3.

Solution:

The vertices and propagators are:

$$= if$$

$$\xrightarrow{\quad k \quad} \;\; = \;\; \frac{-i}{-k^2 + \mu^2 - i\varepsilon}$$

$$\xrightarrow{\quad p \quad} \;\; = \;\; \frac{-i(\gamma^\mu p_\mu + m)}{-p^2 + m^2 - i\varepsilon}$$

Reading the diagram from left to right gives:

$$(\text{Fig. II.5.1}) = \bar{u}(p,s) \int \frac{d^4k}{(2\pi)^4} \, (if) \left(\frac{-i}{-k^2 + \mu^2 - i\epsilon} \right) \left(\frac{-i\,(\slashed{k}+\slashed{p}+ m)}{-(k+p)^2 + m^2 - i\epsilon} \right) (if) \; u(p,s)$$

$$= f^2 \, \bar{u}(p,s) \left[\int \frac{d^4k}{(2\pi)^4} \, \frac{1}{k^2 - \mu^2} \, \frac{(\slashed{k}+\slashed{p}+ m)}{(k+p)^2 - m^2} \right] u(p,s)$$
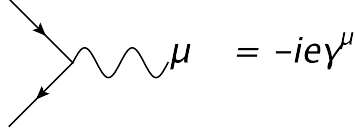
Compare this with page 180 in the book for verification.

2. Applying the Feynman rules for the vector theory (22) show that the amplitude for the diagram in figure II.5.3 is given by
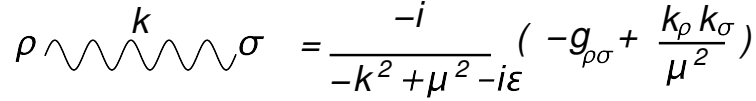
$$(ie)^2 i^2 \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2 - \mu^2} \left(\frac{k_\mu k_\nu}{\mu^2} - g_{\mu\nu} \right) \bar{u}(p)\gamma^\nu \frac{\slashed{p} + \slashed{k} + m}{(p+k)^2 - m^2} \gamma^\mu u(p) \; . \qquad (26)$$

Solution:

The vertices and propagators are:

$$\mu \quad = -ie\gamma^\mu$$

$$\rho \sim\sim\sim^{k}\sim\sim\sim \sigma \quad = \frac{-i}{-k^2 + \mu^2 - i\varepsilon} \left(-g_{\rho\sigma} + \frac{k_\rho k_\sigma}{\mu^2} \right)$$

The fermion propagator is the same as for the previous question. Again, reading the diagram from left to right gives:

(Fig. II.5.3) $=$

$$\bar{u}(p,s) \int \frac{d^4k}{(2\pi)^4} (-ie\gamma^\mu) \left[\frac{-i}{-k^2 + \mu^2 - i\epsilon} \left(-g_{\mu\nu} + \frac{k_\mu k_\nu}{\mu^2} \right) \right] \left[\frac{-i(\slashed{k} + \slashed{p} + m)}{-(k+p)^2 + m^2 - i\epsilon} \right] (-ie\gamma^\nu) \, u(p,s)$$

$$= e^2 \int \frac{d^4k}{(2\pi)^4} \frac{1}{-k^2 + \mu^2} \left(-g_{\mu\nu} + \frac{k_\mu k_\nu}{\mu^2} \right) \bar{u}(p,s)\gamma^\mu \frac{\slashed{k} + \slashed{p} + m}{-(k+p)^2 + m^2} \gamma^\nu u(p,s)$$

$$= e^2 \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2 - \mu^2} \left(-g_{\mu\nu} + \frac{k_\mu k_\nu}{\mu^2} \right) \bar{u}(p,s)\gamma^\nu \frac{\slashed{k} + \slashed{p} + m}{(k+p)^2 - m^2} \gamma^\mu u(p,s)$$

II.6   Scattering and Gauge Invariance

1. Show that the differential cross section for a relativistic electron scattering in a Coulomb potential is given by

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4\vec{p}^2 v^2 \sin^4(\theta/2)} (1 - v^2 \sin^2(\theta/2))$$

Solution:

The amplitude is given in equation (4) on p. 133 as

$$\mathcal{M} = -i \frac{4\pi\alpha}{k^2} \bar{u}'\gamma^0 u$$

where $4\pi\alpha \equiv e^2$, and we have abbreviated $u' \equiv u(P,S)$ and $u \equiv u(p,s)$. Taking the magnitude squared gives

$$|\mathcal{M}|^2 = \left(\frac{4\pi\alpha}{k^2} \right)^2 (\bar{u}'\gamma^0 u)(\bar{u}\gamma^0 u') = \left(\frac{4\pi\alpha}{k^2} \right)^2 \mathrm{tr}\left[\gamma^0(u\bar{u})\gamma^0(u'\bar{u}') \right]$$

57

Summing over spins as $\sum_s u\bar{u} = \frac{1}{2m}(\not{p}+m)$ gives

$$\mathbb{M}^2 \equiv \frac{1}{2}\sum_{s,S}|\mathcal{M}|^2 = \frac{1}{8m^2}\left(\frac{4\pi\alpha}{k^2}\right)^2 \mathrm{tr}\left[\gamma^0(\not{p}+m)\gamma^0(\not{P}+m)\right]$$

The trace over gamma matrices simplifies as

$$
\begin{aligned}
\mathrm{tr}\left[\gamma^0(\not{p}+m)\gamma^0(\not{P}+m)\right] &= \mathrm{tr}(\gamma^0\not{p}\gamma^0\not{P}+m^2 I)\\
&= \mathrm{tr}(-\not{p}\not{P}+2p^0\gamma^0\not{P}+m^2 I)\\
&= 4(-p_\mu P^\mu + 2p^0 P^0 + m^2)\\
&= 4(p^0 P^0 + \vec{p}\cdot\vec{P}+m^2)
\end{aligned}
$$

Therefore the spin-summed amplitude is

$$\mathbb{M}^2 = \frac{1}{2m^2}\left(\frac{4\pi\alpha}{k^2}\right)^2 (p^0 P^0 + \vec{p}\cdot\vec{P}+m^2)$$

where $p^0 = \sqrt{|\vec{p}|^2+m^2}$ and $P^0 = \sqrt{|\vec{P}|^2+m^2}$, and the scattering angle θ is defined as $\vec{p}\cdot\vec{P} = |\vec{p}||\vec{P}|\cos\theta$. The differential scattering cross section is

$$d\sigma = \frac{1}{\gamma|\vec{v}|}\frac{d^3 P}{(2\pi)^3(P^0/m)}\,2\pi\,\delta(p^0 - P^0)\,\mathbb{M}^2$$

where the relativistic factor $\gamma \equiv 1/\sqrt{1-|\vec{v}|^2}$ is included in the denominator since we calculate in the lab frame. The energy conserving delta function implies $|\vec{p}| = |\vec{P}|$, as expected: only the direction of the 3-momentum should change while the magnitude stays fixed. Choosing not to integrate over the solid angle but integrating over $|\vec{P}|$, we get

$$
\begin{aligned}
\frac{d\sigma}{d\Omega} &= \frac{m}{2(2\pi)^2\gamma|\vec{v}|}\left(\frac{|\vec{p}|^2}{\sqrt{|\vec{p}|^2+m^2}}\right)\left(\frac{\sqrt{|\vec{p}|^2+m^2}}{|\vec{p}|}\right)\frac{1}{m^2}\left(\frac{4\pi\alpha}{k^2}\right)^2\left[(|\vec{p}|^2+m^2)+|\vec{p}|^2\cos\theta+m^2\right]\\
&= \frac{|\vec{p}|}{8\pi^2\gamma|\vec{v}|m}\left(\frac{4\pi\alpha}{k^2}\right)^2\left[|\vec{p}|^2(1+\cos\theta)+2m^2\right]
\end{aligned}
$$

Using $k^2 = -4|\vec{p}|^2\sin^2(\theta/2)$ from p. 133 along with $\cos\theta = 1 - 2\sin^2(\theta/2)$, we get

$$
\begin{aligned}
\frac{d\sigma}{d\Omega} &= \frac{|\vec{p}|}{8\pi^2\gamma|\vec{v}|m}\left(\frac{\pi\alpha}{|\vec{p}|^2\sin^2(\theta/2)}\right)^2 2\left[|\vec{p}|^2(1-\sin^2(\theta/2))+m^2\right]\\
&= \frac{\alpha^2}{4\gamma|\vec{v}||\vec{p}|^2\sin^4(\theta/2)}\frac{1}{m|\vec{p}|}\left[|\vec{p}|^2+m^2-|\vec{p}|^2\sin^2(\theta/2)\right]
\end{aligned}
$$

The relativistic 3-momentum is $\vec{p} = \gamma m\vec{v}$, with $\gamma \equiv 1/\sqrt{1-|\vec{v}|^2}$ as before. From now on, since there are no more 4-vectors, let $v \equiv |\vec{v}|$ and $p \equiv |\vec{p}|$. The quantity in square brackets

simplifies as

$$p^2 + m^2 - p^2 \sin^2(\theta/2) = \gamma^2 m^2 v^2 + m^2 - \gamma^2 m^2 v^2 \sin^2(\theta/2)$$
$$= m^2 \left[\frac{v^2}{1 - v^2} + 1 - \frac{v^2}{1 - v^2} \sin^2(\theta/2) \right]$$
$$= \frac{m^2}{1 - v^2} \left[1 - v^2 \sin^2(\theta/2) \right]$$
$$= \gamma^2 m^2 [1 - v^2 \sin^2(\theta/2)]$$

The cross section is therefore

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4\gamma v p^2 \sin^4(\theta/2)} \frac{1}{m(\gamma m v)} \gamma^2 m^2 [1 - v^2 \sin^2(\theta/2)]$$
$$= \frac{\alpha^2}{4v^2 p^2 \sin^4(\theta/2)} [1 - v^2 \sin^2(\theta/2)] \ .$$

2. To order e^2 the amplitude for positron scattering off a proton is just minus the amplitude (3) for electron scattering off a proton. Thus, somewhat counterintuitively, the differential cross sections for positron scattering off a proton and for electron scattering off a proton are the same to this order. Show that to the next order this is no longer true.

Solution:

Consider the 1-loop corrections to the propagator of the exchanged photon. These will all contribute with the same sign to the amplitudes for $e^+ p \to e^+ p$ and $e^- p \to e^- p$. Since the tree level amplitudes differ by a sign, the amplitudes including 1-loop effects are different.

3. Show that the trace of a product of an odd number of gamma matrices vanishes.

Solution:

The gamma matrices can be written in the Weyl basis as $\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu_{\alpha\dot\alpha} \\ \bar\sigma^{\mu\dot\alpha\alpha} & 0 \end{pmatrix}$. In this representation, it is immediately clear that the trace of an odd number of gamma matrices is zero simply because such a quantity cannot be written down. For example,

$$\gamma^\mu \gamma^\nu = \begin{pmatrix} \sigma^\mu_{\alpha\dot\alpha} \bar\sigma^{\nu\dot\alpha\beta} & 0 \\ 0 & \bar\sigma^{\mu\dot\alpha\alpha} \sigma^\nu_{\alpha\dot\beta} \end{pmatrix} \implies \mathrm{tr}(\gamma^\mu \gamma^\nu) = \begin{pmatrix} \mathrm{tr}(\sigma^\mu \bar\sigma^\nu) & 0 \\ 0 & \mathrm{tr}(\bar\sigma^\mu \sigma^\nu) \end{pmatrix}$$

but

$$\gamma^\mu \gamma^\nu \gamma^\rho = \begin{pmatrix} 0 & \sigma^\mu_{\alpha\dot\alpha} \bar\sigma^{\nu\dot\alpha\beta} \sigma^\rho_{\beta\dot\gamma} \\ \bar\sigma^{\mu\dot\alpha\alpha} \sigma^\nu_{\alpha\dot\beta} \bar\sigma^{\rho\dot\beta\gamma} & 0 \end{pmatrix}$$

cannot be traced over, since the spinor indices cannot be contracted.

6. For those who relish long calculations, determine the differential cross section for electron-electron scattering without taking the relativistic limit.

Solution:

For this problem, we use $\sum_s u_s(\vec{p})\bar{u}_s(\vec{p}) = \not{p} + m$ for the spin sum and $\int \frac{d^3p}{(2\pi)^3 2\omega_{\vec{p}}}$ for the measure. The relevant Feynman rules are:

$$\text{electron propagator:} \quad i\,\frac{\not{p}+m}{p^2 - m^2 + i\varepsilon}$$

$$\text{photon propagator:} \quad \frac{-i\,\eta_{\mu\nu}}{p^2 + i\varepsilon} \quad \text{(Feynman gauge)}$$

$$\text{vertex:} \quad = -ie\gamma^{\mu}$$

An incoming electron gets a $u_s(\vec{p})$ and an outgoing electron gets a $\bar{u}_{s'}(\vec{p}')$. We will use the notation $u_i \equiv u_{s_i}(\vec{p}_i)$ and primes for the outgoing particles. There are two diagrams that contribute to $\mathcal{M}$ at tree level:
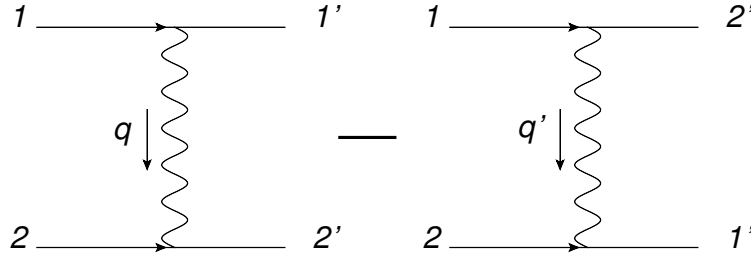
Remembering the relative minus sign between these diagrams gives the amplitude:

$$\mathcal{M} = (\bar{u}_{1'}ie\gamma^{\mu}u_1)\frac{-i\eta_{\mu\nu}}{q^2 + i\varepsilon}(\bar{u}_{2'}ie\gamma^{\nu}u_2) - (\bar{u}_{2'}ie\gamma^{\mu}u_1)\frac{-i\eta_{\mu\nu}}{(q')^2 + i\varepsilon}(\bar{u}_{1'}ie\gamma^{\nu}u_2) + O(e^4)$$

$$= ie^2\left[\frac{(\bar{u}_{1'}\gamma^{\mu}u_1)(\bar{u}_{2'}\gamma_{\mu}u_2)}{(p_1 - p_{1'})^2 + i\varepsilon} - \frac{(\bar{u}_{2'}\gamma^{\mu}u_1)(\bar{u}_{1'}\gamma_{\mu}u_2)}{(p_1 - p_{2'})^2 + i\varepsilon}\right] + O(e^4)$$

Now we need $|\mathcal{M}|^2$. Remember that $\mathcal{M}$ is a Lorentz-scalar (it is just a number), so the conjugation can be carried out as $\mathcal{M}^* = \mathcal{M}^{\dagger}$. However, we are used to working with u and $\bar{u}$ rather than u and $u^{\dagger}$, as required by Lorentz invariance. We have defined a barred spinor as $\bar{\psi} \equiv \psi^{\dagger}\gamma^0$, so for consistency we define a barred matrix as $\overline{M} \equiv \gamma^0 M^{\dagger}\gamma^0$. For instance,

$$\overline{(\gamma^{\mu})} = \gamma^0(\gamma^{\mu})^{\dagger}\gamma^0 = \gamma^0(\gamma^0\gamma^{\mu}\gamma^0)\gamma^0 = \gamma^{\mu} \ .$$

Just as $(AB)^\dagger = B^\dagger A^\dagger$, barring also reverses the order. So to compute the conjugate of the amplitude, we write $|\mathcal{M}|^2 = \mathcal{M}\overline{\mathcal{M}}$:

$$|\mathcal{M}|^2/e^4 = \frac{A}{[(p_1 - p_{1'})^2]^2} + \frac{B}{[(p_1 - p_{2'})^2]^2} - \frac{C + D}{(p_1 - p_{1'})^2(p_1 - p_{2'})^2}$$

$$A = (\bar{u}_{1'}\gamma^\mu u_1)(\bar{u}_{2'}\gamma_\mu u_2)(\bar{u}_1\gamma^\nu u_{1'})(\bar{u}_2\gamma_\nu u_{2'})$$
$$B = (\bar{u}_{2'}\gamma^\mu u_1)(\bar{u}_{1'}\gamma_\mu u_2)(\bar{u}_1\gamma^\nu u_{2'})(\bar{u}_2\gamma_\nu u_{1'})$$
$$C = (\bar{u}_{1'}\gamma^\mu u_1)(\bar{u}_{2'}\gamma_\mu u_2)(\bar{u}_1\gamma^\nu u_{2'})(\bar{u}_2\gamma_\nu u_{1'})$$
$$D = (\bar{u}_1\gamma^\mu u_{1'})(\bar{u}_2\gamma_\mu u_{2'})(\bar{u}_{2'}\gamma^\nu u_1)(\bar{u}_{1'}\gamma_\nu u_2)$$

Each set of parentheses denotes a Lorentz-scalar, meaning that all spinor indices are contracted. To figure out how the Dirac traces work, we write out the indices explicitly and rearrange everything in matrix multiplication order. Denoting $a = 1, 2, 3, 4$ as a Dirac 4-spinor index, we rearrange as follows:

$$A = \bar{u}_{1'}^a(\gamma^\mu u_1)^a \bar{u}_{2'}^b(\gamma_\mu u_2)^b \bar{u}_1^c(\gamma^\nu u_{1'})^c \bar{u}_2^d(\gamma_\nu u_{2'})^d$$
$$= (\gamma^\mu u_1 \bar{u}_1)^{ac}(\gamma^\nu u_{1'}\bar{u}_{1'})^{ca}(\gamma_\mu u_2 \bar{u}_2)^{bd}(\gamma_\nu u_{2'}\bar{u}_{2'})^{db}$$
$$= \mathrm{Tr}\left[(\gamma^\mu u_1\bar{u}_1)(\gamma^\nu u_{1'}\bar{u}_{1'})\right]\ \mathrm{Tr}\left[(\gamma_\mu u_2\bar{u}_2)(\gamma_\nu u_{2'}\bar{u}_{2'})\right]$$

$$B = \bar{u}_{2'}^a(\gamma^\mu u_1)^a \bar{u}_{1'}^b(\gamma_\mu u_2)^b \bar{u}_1^c(\gamma^\nu u_{2'})^c \bar{u}_2^d(\gamma_\nu u_{1'})^d$$
$$= (\gamma^\mu u_1\bar{u}_1)^{ac}(\gamma^\nu u_{2'}\bar{u}_{2'})^{ca}(\gamma_\mu u_2\bar{u}_2)^{bd}(\gamma_\nu u_{1'}\bar{u}_{1'})^{db}$$
$$= \mathrm{Tr}[(\gamma^\mu u_1\bar{u}_1)(\gamma^\nu u_{2'}\bar{u}_{2'})]\,\mathrm{Tr}[(\gamma_\mu u_2\bar{u}_2)(\gamma_\nu u_{1'}\bar{u}_{1'})]$$

$$C = \bar{u}_{1'}^a(\gamma^\mu u_1)^a \bar{u}_{2'}^b(\gamma_\mu u_2)^b \bar{u}_1^c(\gamma^\mu u_{2'})^c \bar{u}_2^d(\gamma_\nu u_{1'})^d$$
$$= (\gamma^\mu u_1\bar{u}_1)^{ac}(\gamma^\nu u_{2'}\bar{u}_{2'})^{cb}(\gamma_\mu u_2\bar{u}_2)^{bd}(\gamma_\nu u_{1'}\bar{u}_{1'})^{da}$$
$$= \mathrm{Tr}[(\gamma^\mu u_1\bar{u}_1)(\gamma^\nu u_{2'}\bar{u}_{2'})(\gamma_\mu u_2\bar{u}_2)(\gamma_\nu u_{1'}\bar{u}_{1'})]$$

$$D = \bar{u}_1^a(\gamma^\mu u_{1'})^a \bar{u}_2^b(\gamma_\mu u_{2'})^b \bar{u}_{2'}^c(\gamma^\nu u_1)^c \bar{u}_{1'}^d(\gamma_\nu u_2)^d$$
$$= (\gamma^\mu u_{1'}\bar{u}_{1'})^{ad}(\gamma_\nu u_2\bar{u}_2)^{db}(\gamma_\mu u_{2'}\bar{u}_{2'})^{bc}(\gamma^\nu u_1\bar{u}_1)^{ca}$$
$$= \mathrm{Tr}[(\gamma^\mu u_{1'}\bar{u}_{1'})(\gamma_\nu u_2\bar{u}_2)(\gamma_\mu u_{2'}\bar{u}_{2'})(\gamma^\nu u_1\bar{u}_1)]$$

Now we would like to sum over all spins and use $\sum_s u_i\bar{u}_i = (\not{p}_i + m)$. To compute the amplitude, we want to average over the initial spins, which gives a factor of $(1/2)$ for each incoming electron. We will multiply by this factor of $(1/2)^2 = 1/4$ later. For now, we simply define the spin-summed version of each of the above pieces:

$$\alpha \equiv \sum_{s_1,s_2,s_{1'},s_{2'}} A \ ,\quad \beta \equiv \sum_{s_1,s_2,s_{1'},s_{2'}} B \ ,\quad \gamma \equiv \sum_{s_1,s_2,s_{1'},s_{2'}} C \ ,\quad \delta \equiv \sum_{s_1,s_2,s_{1'},s_{2'}} D$$

61

We now simplify the above with the following gamma matrix identities:

$$\mathrm{Tr}[\slashed{a}\slashed{b}] = 4\,(ab)$$

$$\mathrm{Tr}[\slashed{a}\slashed{b}\slashed{c}\slashed{d}] = 4\,[(ab)(cd) + (ad)(bc) - (ac)(bd)]$$

$$\mathrm{Tr}[\text{odd number of }\gamma\text{s}] = 0$$

$$\gamma_\mu \gamma^\mu = 4 = \eta^{\mu\nu}\eta_{\mu\nu}$$

$$\gamma_\mu \slashed{a} \gamma^\mu = -2\slashed{a}$$

$$\gamma_\mu \slashed{a}\slashed{b} \gamma^\mu = 4(ab)$$

$$\gamma_\mu \slashed{a}\slashed{b}\slashed{c} \gamma^\mu = -2\slashed{c}\slashed{b}\slashed{a}$$

The notation $(ab) = a_\mu b^\mu$ is used above. It is also useful to define a further condensed notation: $(ij) \equiv p_{i\,\mu} p_j^\mu$. We will only use the latter notation when there is no risk of confusion. Let us now simplify the first term:

$$
\begin{aligned}
\alpha &= \mathrm{Tr}[\gamma^\mu (\slashed{p}_1 + m)\gamma^\nu (\slashed{p}_{1'} + m)]\mathrm{Tr}[\gamma_\mu (\slashed{p}_2 + m)\gamma_\nu (\slashed{p}_{2'} + m)] \\
&= \mathrm{Tr}[(\gamma^\mu \slashed{p}_1 + m\gamma^\mu)(\gamma^\nu \slashed{p}_{1'} + m\gamma^\nu)]\mathrm{Tr}[(\gamma_\mu \slashed{p}_2 + m\gamma_\mu)(\gamma_\nu \slashed{p}_{2'} + m\gamma_\nu)] \\
&= \mathrm{Tr}[\gamma^\mu \slashed{p}_1 \gamma^\nu \slashed{p}_{1'} + m^2 \gamma^\mu \gamma^\nu]\mathrm{Tr}[\gamma_\mu \slashed{p}_2 \gamma_\nu \slashed{p}_{2'} + m^2 \gamma_\mu \gamma_\nu] \\
&= 4[(p_{1'}^\mu p_1^\nu + p_{1'}^\nu p_1^\mu) - (p_1 p_{1'})\eta^{\mu\nu} + 4m^2 \eta^{\mu\nu}]4[(p_{2\mu}p_{2'\nu} + p_{2\nu}p_{2'\mu}) - (p_2 p_{2'})\eta_{\mu\nu} + 4m^2 \eta_{\mu\nu}]
\end{aligned}
$$

The parentheses on the first term inside each [...] is there just to show that the term is symmetric in μ and ν, and it is thus useful to organize the calculation using this grouping. Continuing:

$$
\begin{aligned}
\alpha &= 16[(p_{1'}^\mu p_1^\nu + p_{1'}^\nu p_1^\mu)(p_{2\mu}p_{2'\nu} + p_{2\nu}p_{2'\mu}) - 2(p_2 p_{2'})(p_{1'}p_1) + 4m^2 \times 2(p_1 p_{1'}) \\
&\quad - 2(p_1 p_{1'})(p_2 p_{2'}) + (p_1 p_{1'})(p_2 p_{2'})\underbrace{\eta^{\mu\nu}\eta_{\mu\nu}}_{=+4} - 4m^2 (p_1 p_{1'})\eta^{\mu\nu}\eta_{\mu\nu} + 4m^2 \times 2(p_2 p_{2'}) \\
&\quad - 4m^2 (p_2 p_{2'})\eta^{\mu\nu}\eta_{\mu\nu} + 16m^4 \eta^{\mu\nu}\eta_{\mu\nu}] \\[2mm]
&= 16[(1'2)(12') + (1'2')(12) + (1'2')(12) + (1'2)(12') - 2(22')(1'1) + 8m^2(1'1) \\
&\quad - 2(11')(22') + 4(11')(22') - 16m^2(11') + 8m^2(22') - 16m^2(22') + 64m^4] \\[2mm]
&= 16[2(1'2)(12') + 2(1'2')(12) - 8m^2(11') - 8m^2(22') + 64m^4] \\[2mm]
&= 32[(1'2)(12') + (12)(1'2') - 4m^2(11') - 4m^2(22') + 32m^4]
\end{aligned}
$$

A quick check from dimensional analysis: Each (ij) has dimensions of m^2, so each term has dimensions of m^4, so the result is at least dimensionally correct. The next term is:

$$\beta = \text{Tr}[\gamma^\mu (\not{p}_1 + m)\gamma^\nu (\not{p}_{2'} + m)]\text{Tr}[\gamma_\mu (\not{p}_2 + m)\gamma_\nu (\not{p}_{1'} + m)]$$

$$= \text{Tr}[(\gamma^\mu \not{p}_1 + m\gamma^\mu)(\gamma^\nu \not{p}_{2'} + m\gamma^\nu)]\text{Tr}[(\gamma_\mu \not{p}_2 + m\gamma_\mu)(\gamma_\nu \not{p}_{1'} + m\gamma_\nu)]$$

$$= \text{Tr}[\gamma^\mu \not{p}_1 \gamma^\nu \not{p}_{2'} + m^2 \gamma^\mu \gamma^\nu]\text{Tr}[\gamma_\mu \not{p}_2 \gamma_\nu \not{p}_{1'} + m^2 \gamma_\mu \gamma_\nu]$$

$$= 16[(p_1^\mu p_{2'}^\nu + p_1^\nu p_{2'}^\mu) - (p_1 p_{2'})\eta^{\mu\nu} + 4m^2 \eta^{\mu\nu}][(p_{2\mu} p_{1'\nu} + p_{2\nu} p_{1'\mu}) - (p_2 p_{1'})\eta_{\mu\nu} + 4m^2 \eta_{\mu\nu}]$$

$$= 16[(p_1^\mu p_{2'}^\nu + p_1^\nu p_{2'}^\mu)(p_{2\mu} p_{1'\nu} + p_{2\nu} p_{1'\mu}) - 2(12')(21') + 4m^2 \times 2(12')$$
$$- 2(12')(21') + 4(12')(21') - 4m^2 (12') \times 4 + 8m^2 (21') - 4m^2 (21') \times 4 + 16m^4 \times 4]$$

$$= 16[2(12)(1'2') + 2(11')(22') - 2(12')(1'2) + 8m^2 (12')$$
$$- 2(12')(1'2) + 4(12')(1'2) - 16m^2 (12') + 8m^2 (1'2) - 16m^2 (1'2) + 64m^4]$$

$$= 16[2(12)(1'2') + 2(11')(22') - 8m^2 (1'2) - 8m^2 (12') + 64m^4]$$

$$= 32[(12)(1'2') + (11')(22') - 4m^2 (1'2) - 4m^2 (12') + 32m^4]$$

Now for the longer terms, the first of which is:

$$\gamma = \text{Tr}[\gamma^\mu(\not{p}_1 + m)\gamma^\nu(\not{p}_{2'} + m)\gamma_\mu(\not{p}_2 + m)\gamma_\nu(\not{p}_{1'} + m)]$$
$$= \text{Tr}[(\gamma^\mu \not{p}_1 + m\gamma^\mu)(\gamma^\nu \not{p}_{2'} + m\gamma^\nu)(\gamma_\mu \not{p}_2 + m\gamma_\mu)(\gamma_\nu \not{p}_{1'} + m\gamma_\nu)]$$
$$= \text{Tr}[(\gamma^\mu \not{p}_1\gamma^\nu \not{p}_{2'} + m\gamma^\mu \not{p}_1\gamma^\nu + m\gamma^\mu\gamma^\nu \not{p}_{2'} + m^2\gamma^\mu\gamma^\nu)$$
$$\times (\gamma_\mu \not{p}_2\gamma_\nu \not{p}_{1'} + m\gamma_\mu \not{p}_2\gamma_\nu + m\gamma_\mu\gamma_\nu \not{p}_{1'} + m^2\gamma_\mu\gamma_\nu)]$$

$$= \text{Tr}[\underbrace{\gamma^\mu \not{p}_1\gamma^\nu \not{p}_{2'}\gamma_\mu}_{-2\not{p}_{2'}\gamma^\nu \not{p}_1} \not{p}_2\gamma_\nu \not{p}_{1'} + 0 + 0 + m^2 \underbrace{\gamma^\mu \not{p}_1\gamma^\nu \not{p}_{2'}\gamma_\mu}_{-2\not{p}_{2'}\gamma^\nu \not{p}_1} \gamma_\nu$$
$$+ 0 + m^2 \underbrace{\gamma^\mu \not{p}_1\gamma^\nu\gamma_\mu}_{4p_1^\nu} \not{p}_2\gamma_\nu + m^2 \underbrace{\gamma^\mu \not{p}_1\gamma^\nu\gamma_\mu}_{4p_1^\nu} \gamma_\nu \not{p}_{1'} + 0$$
$$+ 0 + m^2 \underbrace{\gamma^\mu\gamma^\nu \not{p}_{2'}\gamma_\mu}_{4p_{2'}^\nu} \not{p}_2\gamma_\nu + m^2 \underbrace{\gamma^\mu\gamma^\nu \not{p}_{2'}\gamma_\mu}_{4p_{2'}^\nu} \gamma_\nu \not{p}_{1'} + 0$$
$$+ m^2\gamma^\mu \underbrace{\gamma^\nu\gamma_\mu \not{p}_2\gamma_\nu}_{4p_{2\mu}} \not{p}_{1'} + 0 + 0 + m^4 \underbrace{\gamma^\mu\gamma^\nu\gamma_\mu}_{-2\gamma^\nu} \gamma_\nu]$$

$$= \text{Tr}[-2\not{p}_{2'} \underbrace{\gamma^\nu \not{p}_1 \not{p}_2\gamma_\nu}_{4(p_1 p_2)} \not{p}_{1'} - 2m^2\not{p}_{2'} \underbrace{\gamma^\nu \not{p}_1\gamma_\nu}_{-2\not{p}_1} + 4m^2\not{p}_2 \not{p}_1$$
$$+ 4m^2\not{p}_1 \not{p}_{1'} + 4m^2\not{p}_{2'} \not{p}_{1'} + 4m^2\not{p}_2 \not{p}_{1'} - 2m^4 \underbrace{\gamma^\nu\gamma_\nu}_{4}]$$

$$= -8(p_1 p_2)\text{Tr}[\not{p}_{2'} \not{p}_{1'}] + 4m^2\text{Tr}[\not{p}_{2'} \not{p}_1 + \not{p}_2 \not{p}_1 + \not{p}_1 \not{p}_{1'} + \not{p}_2 \not{p}_{2'} + \not{p}_{2'} \not{p}_{1'} + \not{p}_2 \not{p}_{1'}] - 8m^4\text{Tr}[\mathbb{I}_4]$$

$$= -32(p_1 p_2)(p_{1'} p_{2'}) + 16m^2[(p_{2'} p_1) + (p_2 p_1) + (p_1 p_{1'}) + (p_2 p_{2'}) + (p_{1'} p_{2'}) + (p_2 p_{1'})] - 32m^4$$

One more:

$$\delta = \text{Tr}[\gamma^\mu(\not{p}_{1'} + m)\gamma_\nu(\not{p}_2 + m)\gamma_\mu(\not{p}_{2'} + m)\gamma^\nu(\not{p}_1 + m)]$$

$$= \text{Tr}[(\gamma^\mu \not{p}_{1'} + m\gamma^\mu)(\gamma_\nu \not{p}_2 + m\gamma_\nu)(\gamma_\mu \not{p}_{2'} + m\gamma_\mu)(\gamma^\nu \not{p}_1 + m\gamma^\nu)]$$

$$= \text{Tr}[(\gamma^\mu \not{p}_{1'}\gamma_\nu \not{p}_2 + m\gamma^\mu \not{p}_{1'}\gamma_\nu + m\gamma^\mu \gamma_\nu \not{p}_2 + m^2\gamma^\mu \gamma_\nu)$$
$$\times (\gamma_\mu \not{p}_{2'}\gamma^\nu \not{p}_1 + m\gamma_\mu \not{p}_{2'}\gamma^\nu + m\gamma_\mu \gamma^\nu \not{p}_1 + m^2\gamma_\mu \gamma^\nu)]$$

$$= \text{Tr}[\underbrace{\gamma^\mu \not{p}_{1'}\gamma_\nu \not{p}_2 \gamma_\mu}_{-2\not{p}_2 \gamma_\nu \not{p}_{1'}} \not{p}_{2'}\gamma^\nu \not{p}_1 + 0 + 0 + m^2 \underbrace{\gamma^\mu \not{p}_{1'}\gamma_\nu \not{p}_2 \gamma_\mu}_{-2\not{p}_2 \gamma_\nu \not{p}_{1'}} \gamma^\nu$$
$$+ 0 + m^2 \underbrace{\gamma^\mu \not{p}_{1'}\gamma_\nu \gamma_\mu}_{4p_{1'\nu}} \not{p}_{2'}\gamma^\nu + m^2 \underbrace{\gamma^\mu \not{p}_{1'}\gamma_\nu \gamma_\mu}_{4p_{1'\nu}} \gamma^\nu \not{p}_1 + 0$$
$$+ 0 + m^2 \underbrace{\gamma^\mu \gamma_\nu \not{p}_2 \gamma_\mu}_{4p_{2\nu}} \not{p}_{2'}\gamma^\nu + m^2 \underbrace{\gamma^\mu \gamma_\nu \not{p}_2 \gamma_\mu}_{4p_{2\nu}} \gamma^\nu \not{p}_1 + 0$$
$$+ m^2\gamma^\mu \underbrace{\gamma_\nu \gamma_\mu \not{p}_{2'}\gamma^\nu}_{4p_{2'\mu}} \not{p}_1 + 0 + 0 + m^4 \underbrace{\gamma^\mu \gamma_\nu \gamma_\mu}_{-2\gamma_\nu} \gamma^\nu]$$
$$= \text{Tr}[-2\not{p}_2 \underbrace{\gamma_\nu \not{p}_{1'} \not{p}_{2'}\gamma^\nu}_{4(p_{1'}p_{2'})} \not{p}_1 - 2m^2 \not{p}_2 \underbrace{\gamma_\nu \not{p}_{1'}\gamma^\nu}_{-2\not{p}_{1'}}$$
$$+ 4m^2(\not{p}_{2'}\not{p}_{1'} + \not{p}_{1'}\not{p}_1 + \not{p}_{2'}\not{p}_2 + \not{p}_2\not{p}_1 + \not{p}_{2'}\not{p}_1) - 2m^4 \underbrace{\gamma_\nu \gamma^\nu}_{4}]$$

$$= -8(p_{1'}p_{2'})\text{Tr}[\not{p}_2\not{p}_1] + 4m^2\text{Tr}[\not{p}_{2'}\not{p}_{1'} + \not{p}_{1'}\not{p}_1 + \not{p}_{2'}\not{p}_2 + \not{p}_2\not{p}_1 + \not{p}_{2'}\not{p}_1] - 8m^4\text{Tr}[\mathbb{I}_4]$$

$$= -32(p_{1'}p_{2'})(p_1p_2) + 16m^2[(p_{1'}p_2) + (p_{1'}p_{2'}) + (p_1p_{1'}) + (p_2p_{2'}) + (p_1p_2) + (p_1p_{2'})] - 32m^4$$

We see that this equals the previous term: $\gamma = \delta$. Now we have all the terms we need to compute the amplitude-squared. Define the amplitude-squared after averaging over initial spins and summing over final spins: $\mathbb{M}^2 \equiv \frac{1}{4}\sum_{\{\text{all spins}\}} |\mathcal{M}|^2$. We have:

$$\mathbb{M}^2 = \frac{e^4}{4}\left[\frac{\alpha}{[(p_1 - p_{1'})^2]^2} + \frac{\beta}{[(p_1 - p_{2'})^2]^2} - \frac{2\delta}{(p_1 - p_{1'})^2(p_1 - p_{2'})^2}\right]$$

$$\alpha = 32[(1'2)(12') + (12)(1'2') - 4m^2\{(11') + (22')\} + 32m^4]$$
$$\beta = 32[(12)(1'2') + (11')(22') - 4m^2\{(1'2) + (1'2)\} + 32m^4]$$
$$\delta = -16[2(12)(1'2') - m^2\{(12) + (1'2) + (1'2') + (11') + (22') + (12')\} + 2m^4]$$

where $(ij) \equiv p_{i\,\mu}p_j^\mu$

Until now, we have not used any information about the actual kinematics of the scattering process. Remembering that all external particles are on-shell with equal masses, we know $p_i^2 = m^2$, where m is the mass of the electron. We now define the Mandelstam variables s, t, u such that $s + t + u = \sum_i m_i^2 = 4m^2$:

$$s \equiv (p_1 + p_2)^2 = (p_1^2 + p_2^2 + 2(12)) = 2m^2 + 2(12) \implies (12) = \frac{1}{2}(s - 2m^2)$$
$$= (p_{1'} + p_{2'})^2 \implies (1'2') = (12)$$

$$t \equiv (p_1 - p_{1'})^2 = (p_1^2 + p_{1'}^2 - 2(11')) = 2m^2 - 2(11') \implies (11') = -\frac{1}{2}(t - 2m^2)$$
$$= (p_2 - p_{2'})^2 \implies (22') = (11')$$

$$u \equiv (p_1 - p_{2'})^2 = (p_1^2 + p_{2'}^2 - 2(12')) = 2m^2 - 2(12') \implies (12') = -\frac{1}{2}(u - 2m^2)$$
$$= (p_2 - p_{1'})^2 \implies (21') = (12')$$

Therefore, the pieces of the amplitude-squared are:

$$\alpha = 32[(1'2)^2 + (12)^2 - 8m^2(11') + 32m^4]$$
$$\beta = 32[(12)^2 + (11')^2 - 8m^2(1'2) + 32m^4]$$
$$\delta = -16[2(12)^2 - 2m^2\{(12') + (12) + (11')\} + 2m^4]$$

Now that we have gotten this far, we will drop the mass of the electron to check with the results in the text. Setting $m = 0$ simplifies the dot products: $(12) = +s/2$, $(11') = -t/2$, $(12') = -u/2$, where now $s + t + u = 0$. The pieces of the amplitude-squared are now:

$$\alpha = 32[\frac{1}{4}u^2 + \frac{1}{4}s^2] = 8(s^2 + u^2)$$
$$\beta = 32[\frac{1}{4}s^2 + \frac{1}{4}t^2] = 8(s^2 + t^2)$$
$$\delta = -16[\frac{1}{2}s^2] = -8s^2$$

The amplitude-squared is therefore:

$$\mathbb{M}^2 = \frac{e^4}{4}\left[\frac{\alpha}{t^2} + \frac{\beta}{u^2} - \frac{2\delta}{tu}\right]$$
$$= \frac{e^4}{4}\left[\frac{8(s^2 + u^2)}{t^2} + \frac{8(s^2 + t^2)}{u^2} - \frac{2(-8s^2)}{tu}\right]$$
$$= 2e^4\left[\frac{s^2 + u^2}{t^2} + \frac{s^2 + t^2}{u^2} + \frac{2s^2}{tu}\right]$$

66

From the definition of s, we see that $s = E_{CM}^2$, where E_{CM} is the total energy in the center-of-mass frame. We now calculate the differential cross section, whose formula is given in Appendix C. The differential cross section in the center-of-mass frame is:

$$d\sigma_{CM} = \frac{1}{4|\vec{p}_1|_{CM}\sqrt{s}} \mathbb{M}^2 \frac{d^3p_{1'}}{(2\pi)^3 2p_{1'}^0} \frac{d^3p_{2'}}{(2\pi)^3 2p_{2'}^0} (2\pi)^4 \delta^4(p_1 + p_2 - p_{1'} - p_{2'})$$

where $p_i^0 = \sqrt{|\vec{p}_i|^2 + m^2} = |\vec{p}_i|$ for $m = 0$. For 4 identical particles with zero mass, the prefactor can be simplified:

$$|\vec{p}_1|_{CM}\sqrt{s} = \frac{1}{2}\sqrt{s^2 - 2(m_1^2 + m_2^2)s + (m_1^2 - m_2^2)^2} = \frac{1}{2}s$$

Now, the 4-dimensional delta function is a product of a one dimensional energy-conserving delta function and a 3-dimensional momentum-conserving delta function. As discussed, the total energy of the incoming electrons in the center-of-mass frame is $p_1^0 + p_2^0 = \sqrt{s}$. Furthermore, the center-of-mass frame is defined such that $\vec{p}_1 + \vec{p}_2 = \vec{0}$, so the differential cross section is:

$$d\sigma_{CM} = \frac{1}{2s} \mathbb{M}^2 \frac{d^3p_{1'}}{(2\pi)^3 2p_{1'}^0} \frac{d^3p_{2'}}{(2\pi)^3 2p_{2'}^0} (2\pi)^4 \delta^4(p_1 + p_2 - p_{1'} - p_{2'})$$

$$= \frac{\mathbb{M}^2}{2^3(2\pi)^2 s} \frac{d^3p_{1'}}{|\vec{p}_{1'}|} \frac{d^3p_{2'}}{|\vec{p}_{2'}|} \delta(\sqrt{s} - |\vec{p}_{1'}| - |\vec{p}_{2'}|) \delta^3(\vec{p}_{1'} + \vec{p}_{2'})$$

We see that the 3-dimensional delta function simply sets $\vec{p}_{2'} = -\vec{p}_{1'}$:

$$\int d^3p_{2'} f(\vec{p}_{1'}, \vec{p}_{2'}) \delta^3(\vec{p}_{1'} + \vec{p}_{2'}) = f(\vec{p}_{1'}, -\vec{p}_{1'})$$

Therefore, integrating over $\vec{p}_{2'}$ along with the fact that $\delta(2x) = \frac{1}{2}\delta(x)$ gives:

$$d\sigma_{CM} = \frac{\mathbb{M}^2}{2^4(2\pi)^2 s} \frac{d^3p_{1'}}{|\vec{p}_{1'}|^2} \delta(\sqrt{s} - |\vec{p}_{1'}|)$$

Now, using spherical coordinates for $\vec{p}_{1'}$ gives $d^3p_{1'} = d\Omega\, dp\, p^2$, where Ω is the solid angle over which you are not asked to integrate, and $p \equiv |\vec{p}_{1'}|$. The differential cross section per solid angle is now a 1-dimensional integral:

$$\left(\frac{d\sigma}{d\Omega}\right)_{CM} = \frac{1}{2^4(2\pi)^2 s} \int_0^\infty dp\, \mathbb{M}^2 \delta\left(\frac{\sqrt{s}}{2} - p\right)$$

The delta function just sets $p = \sqrt{s}/2$, so all we have to do now is figure out how these integrals over delta functions affected $\mathbb{M}^2$. We can do this by looking at the Mandelstam

variables:

$$t = (p_1 - p_{1'})^2 = -2p_{1\mu}p_{1'}^\mu = 2(-p_1^0 p_{1'}^0 + \vec{p}_1 \cdot \vec{p}_{1'})$$
$$= 2(-|\vec{p}_1||\vec{p}_{1'}| + |\vec{p}_1||\vec{p}_{1'}| \cos\theta)$$
$$= -2|\vec{p}_1||\vec{p}_{1'}|(1 - \cos\theta) = -\sqrt{s}|\vec{p}_{1'}|(1 - \cos\theta)$$
$$= -\sqrt{s}\left(\frac{\sqrt{s}}{2}\right)(1 - \cos\theta) = -\frac{s}{2}(1 - \cos\theta)$$
$$= -s \sin^2(\theta/2)$$

Using $s + t + u = 0$, the other kinematic variable is

$$u = -s - t = -s + \frac{s}{2}(1 - \cos\theta) = -\frac{s}{2} + s\cos\theta = -\frac{s}{2}(1 + \cos\theta) = -s \cos^2(\theta/2)$$

The amplitude-squared is therefore:

$$\mathbb{M}^2 = 2e^4 \left[\frac{s^2 + u^2}{t^2} + \frac{s^2 + t^2}{u^2} + \frac{2s^2}{tu}\right]$$
$$= 2e^4 \underbrace{\left[\frac{1 + \cos^4(\theta/2)}{\sin^4(\theta/2)} + \frac{1 + \sin^4(\theta/2)}{\cos^4(\theta/2)} + \frac{2}{\cos^2(\theta/2)\sin^2(\theta/2)}\right]}_{\equiv f(\theta)}$$

Putting it all together:

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}} = \frac{\mathbb{M}^2}{2^4(2\pi)^2 s} = \frac{2e^4 f(\theta)}{2^4(2\pi)^2 s} = \left(\frac{e^4}{2^5 \pi^2}\right)\frac{f(\theta)}{s} = \left(\frac{e^2}{4\pi}\right)^2 \frac{f(\theta)}{2E_{\text{CM}}^2}$$

Finally, $E_{\text{CM}} = E_1 + E_2 = 2E \implies E_{\text{CM}}^2 = 4E$, so we get the result:

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{CM}} = \left(\frac{e^2}{4\pi}\right)^2 \frac{f(\theta)}{8E^2}$$

II.7   Diagrammatic Proof of Gauge Invariance

1. Extend the proof to cover figure II.7.1c. [Hint: To get oriented, note that figure II.7.1b corresponds to $n = 1$.]

Solution:

We will work out the case $n = 2$, which will make the generalization to arbitrary n clear. There are now $(2+1)! = 6$ diagrams, corresponding to the different ways in which the photon lines can cross. We will compute them in groups of three.

First consider the three diagrams given below:

68
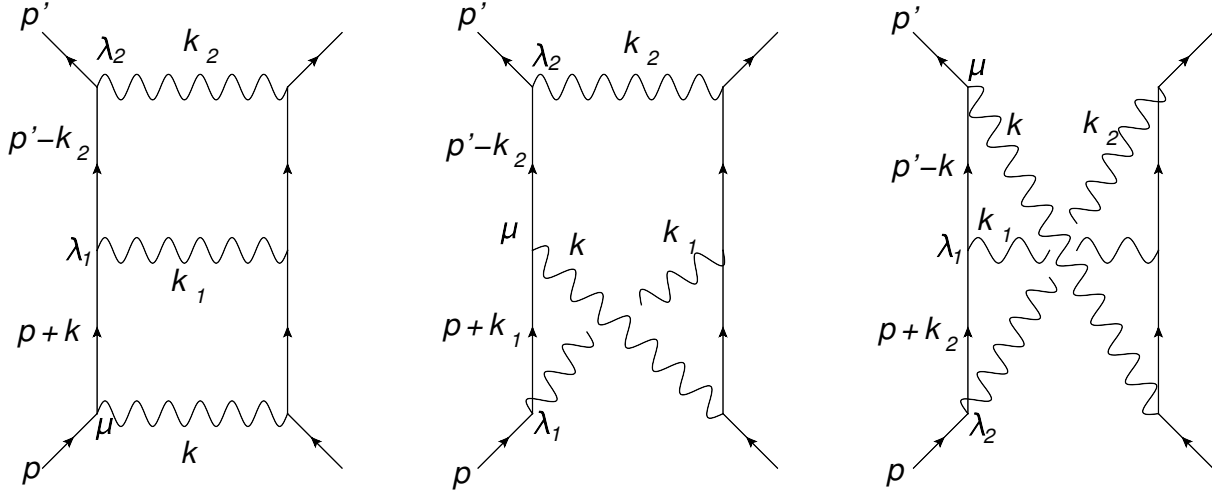
where we choose the convention in which the internal photon momenta always flow towards the left. The three ks are related by $p' - p = k + k_1 + k_2$. The amplitude given by adding (bose statistics) these three diagrams is:

$$\mathcal{M}_1 = \bar{u}' \left[\gamma^{\lambda_2} \frac{1}{\not{p}' - \not{k}_2} \gamma^{\lambda_1} \frac{1}{\not{p} + \not{k}} \not{k} + \gamma^{\lambda_2} \frac{1}{\not{p}' - \not{k}_2} \not{k} \frac{1}{\not{p} + \not{k}_1} \gamma^{\lambda_1} + \not{k} \frac{1}{\not{p}' - \not{k}} \gamma^{\lambda_1} \frac{1}{\not{p} + \not{k}_2} \gamma^{\lambda_2} \right] u \ .$$

Here we have dropped the fermion masses since they play no role in the proof. We have also suppressed all irrelevant factors in the amplitude, retaining only the k_μ term in the photon propagator, as explained in the text.

In the numerator of the first term, write $\not{k} = (\not{p} + \not{k}) - \not{p}$, and in the third term write $\not{k} = -(\not{p}' - \not{k}) + \not{p}'$. The Dirac equation is now $\not{p} u = 0$ and $\not{p}' u' = 0$, so that the $\not{p}$ gives zero acting to the right, and $\not{p}'$ gives zero acting to the left. We have

$$\mathcal{M}_1 = \bar{u}' \left[\gamma^{\lambda_2} \frac{1}{\not{p}' - \not{k}_2} \gamma^{\lambda_1} + \gamma^{\lambda_2} \frac{1}{\not{p}' - \not{k}_2} \not{k} \frac{1}{\not{p} + \not{k}_1} \gamma^{\lambda_1} - \gamma^{\lambda_1} \frac{1}{\not{p} + \not{k}_2} \gamma^{\lambda_2} \right] u$$

$$= \bar{u}' \left[\gamma^{\lambda_2} \frac{1}{\not{p} + \not{k} + \not{k}_1} \gamma^{\lambda_1} + \gamma^{\lambda_2} \frac{1}{\not{p} + \not{k} + \not{k}_1} \not{k} \frac{1}{\not{p} + \not{k}_1} \gamma^{\lambda_1} - \gamma^{\lambda_1} \frac{1}{\not{p} + \not{k}_2} \gamma^{\lambda_2} \right] u$$

where in the second line we have written $p' - k_2 = p + k + k_1$. In the numerator of the middle term, write $k = (p + k_1 + k) - (p + k_1)$ to get:

$$\mathcal{M}_1 = \bar{u}' \left[\gamma^{\lambda_2} \frac{1}{\not{p} + \not{k} + \not{k}_1} \gamma^{\lambda_1} + \gamma^{\lambda_2} \frac{1}{\not{p} + \not{k}_1} \gamma^{\lambda_1} - \gamma^{\lambda_2} \frac{1}{\not{p} + \not{k} + \not{k}_1} \gamma^{\lambda_1} - \gamma^{\lambda_1} \frac{1}{\not{p} + \not{k}_2} \gamma^{\lambda_2} \right] u$$

$$= \bar{u}' \left[\gamma^{\lambda_2} \frac{1}{\not{p} + \not{k}_1} \gamma^{\lambda_1} - \gamma^{\lambda_1} \frac{1}{\not{p} + \not{k}_2} \gamma^{\lambda_2} \right] u$$

where we have canceled the first and third terms from the first line. We see that the remaining two terms do not cancel; indeed, this is one way to discover additional diagrams in case you had forgotten them. That is, it is obvious that there are at least 3 diagrams but it may have escaped your attention that there are actually a total of $3! = 6$ diagrams.

The remaining three diagrams are displayed below:
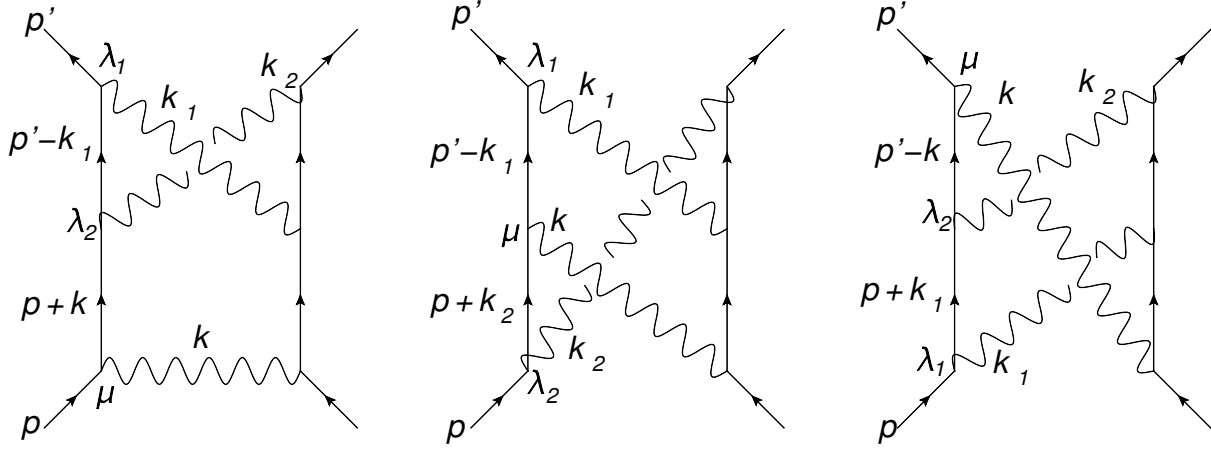
As before, two of the resulting terms will cancel each other and the remaining terms will cancel the leftover terms in $\mathcal{M}_1$. The amplitude from these three new diagrams is:

$$\mathcal{M}_2 = \bar{u}\left[\gamma^{\lambda_1}\frac{1}{\slashed{p}'-\slashed{k}_1}\gamma^{\lambda_2}\frac{1}{\slashed{p}+\slashed{k}}\;\slashed{k} + \gamma^{\lambda_1}\frac{1}{\slashed{p}'-\slashed{k}_1}\;\slashed{k}\frac{1}{\slashed{p}+\slashed{k}_2}\gamma^{\lambda_2} + \slashed{k}\frac{1}{\slashed{p}'-\slashed{k}}\gamma^{\lambda_2}\frac{1}{\slashed{p}+\slashed{k}_1}\gamma^{\lambda_1}\right]u\;.$$

In the numerator of the first term, write $k = (p+k) - p$ and in the third term write $k = -(p'-k)+p'$. Again $\slashed{p}$ acting to the right gives zero, and $\slashed{p}'$ acting to the left gives zero. The amplitude is

$$\mathcal{M}_2 = \bar{u}\left[\gamma^{\lambda_1}\frac{1}{\slashed{p}'-\slashed{k}_1}\gamma^{\lambda_2} + \gamma^{\lambda_1}\frac{1}{\slashed{p}'-\slashed{k}_1}\;\slashed{k}\frac{1}{\slashed{p}+\slashed{k}_2}\gamma^{\lambda_2} - \gamma^{\lambda_2}\frac{1}{\slashed{p}+\slashed{k}_1}\gamma^{\lambda_1}\right]u$$

$$= \bar{u}\left[\gamma^{\lambda_1}\frac{1}{\slashed{p}+\slashed{k}+\slashed{k}_2}\gamma^{\lambda_2} + \gamma^{\lambda_1}\frac{1}{\slashed{p}+\slashed{k}+\slashed{k}_2}\;\slashed{k}\frac{1}{\slashed{p}+\slashed{k}_2}\gamma^{\lambda_2} - \gamma^{\lambda_2}\frac{1}{\slashed{p}+\slashed{k}_1}\gamma^{\lambda_1}\right]u$$

where we have used $p' - k_1 = p + k + k_2$. In the numerator of the middle term, write $k = (p+k+k_2) - (p+k_2)$ to get:

$$\mathcal{M}_2 = \bar{u}\left[\gamma^{\lambda_1}\frac{1}{\slashed{p}+\slashed{k}+\slashed{k}_2}\gamma^{\lambda_2} + \gamma^{\lambda_1}\frac{1}{\slashed{p}+\slashed{k}_2}\gamma^{\lambda_2} - \gamma^{\lambda_1}\frac{1}{\slashed{p}+\slashed{k}+\slashed{k}_2}\gamma^{\lambda_2} - \gamma^{\lambda_2}\frac{1}{\slashed{p}+\slashed{k}_1}\gamma^{\lambda_1}\right]u$$

$$= \bar{u}\left[\gamma^{\lambda_1}\frac{1}{\slashed{p}+\slashed{k}_2}\gamma^{\lambda_2} - \gamma^{\lambda_2}\frac{1}{\slashed{p}+\slashed{k}_1}\gamma^{\lambda_1}\right]u = -\mathcal{M}_1$$

Thus the contribution from the offending term $k_\mu k_\nu/\mu^2$ in the photon propagator contributes a total of $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2 = 0$ to the amplitude. In general, the photon with momentum k can connect to the left-most line in $(n+1)$ ways, which results in $(n+1)!$ different diagrams, whose constituent terms will cancel in pairs.

2. You might have worried whether the shift of integration variable is allowed. Rationalizing the denominators in the first integral

$$\int\frac{d^4p}{(2\pi)^4}\text{tr}\left(\gamma^\nu\frac{1}{\slashed{p}_2-m}\gamma^\sigma\frac{1}{\slashed{p}_1-m}\gamma^\lambda\frac{1}{\slashed{p}-m}\right)$$

70

in (13) and imagining doing the trace, you can convince yourself that this integral is only logarithmically divergent and hence that the shift is allowed. This issue will come up again in chapter IV.7 and we are anticipating a bit here.

Solution:

Rationalizing the denominators gives

$$I \equiv \int \frac{d^4 p}{(2\pi)^4} \text{tr} \left(\gamma^\nu \frac{1}{\not{p}_2 - m} \gamma^\sigma \frac{1}{\not{p}_1 - m} \gamma^\lambda \frac{1}{\not{p} - m} \right) = \int \frac{d^4 p}{(2\pi)^4} \frac{N^{\nu\sigma\lambda}}{(p_2^2 - m^2)(p_1^2 - m^2)(p^2 - m^2)}$$

where the numerator is

$$N^{\nu\sigma\lambda} = \text{tr} \left[\gamma^\nu (\not{p}_2 + m) \gamma^\sigma (\not{p}_1 + m) \gamma^\lambda (\not{p} + m) \right]$$

and $p_1 \equiv p + q_1$ and $p_2 \equiv p + q_2$. The key is that after suitably combining denominators and shifting integration variables, the cubic terms in the numerator will vanish by Lorentz invariance. For large loop momentum p the integral behaves as

$$I \sim \int d^4 p \frac{p^2}{(p^2)^3} \sim \int dp \, p^3 \frac{p^2}{p^6} \sim \int dp \frac{1}{p} \sim \ln p$$

which depends logarithmically on the cutoff.

II.8 Photon-Electron Scattering and Crossing

1. Show that averaging and summing over photon polarizations amounts to replacing the square bracket in (9) by $2[\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2 \theta]$. [Hint: We are working in the transverse gauge.]

Solution:

Let us first derive a general Lorentz-covariant definition of the sum over photon polarizations.

For a circularly polarized[3] photon moving in the $+\hat{z}$ direction, the polarization vectors can be taken as

$$\varepsilon_1^\mu = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ +i \\ 0 \end{pmatrix} , \quad \varepsilon_2^\mu = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -i \\ 0 \end{pmatrix} .$$

Explicitly computing the sum over polarizations gives

$$\sum_{a=1,2} \varepsilon_a^\mu \varepsilon_a^{\nu*} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} .$$

---

[3]We are using circular polarization to show an example with complex polarization vectors.

We would like to write this in terms of Lorentz-covariant notation. The 4-momentum of the photon is $k^\mu = |\vec{k}|(1,0,0,1)^T$, so that

$$\frac{k^\mu k^\nu}{|\vec{k}|^2} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Define the spacelike unit vector in the direction of propagation of the photon: $n^\mu = (0,0,0,1)^T$. Then we have

$$\frac{k^\mu n^\nu + k^\nu n^\mu}{|\vec{k}|} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix}.$$

Furthermore, we can write $|\vec{k}| = -k \cdot n$. Therefore, using the Lorentz-invariant metric $\eta^{\mu\nu} = \text{diag}(+1,-1,-1,-1)$ we can write

$$\sum_{a=1,2} \varepsilon_a^\mu(k) \varepsilon_a^\nu(k)^* = -\eta^{\mu\nu} + \frac{k^\mu n^\nu + k^\nu n^\mu}{k \cdot n} + \frac{k^\mu k^\nu}{(k \cdot n)^2}.$$

This is valid for any reference frame.

We may now proceed with the problem. We can use the above result to write:

$$S \equiv \left[\sum_{a=1,2} \varepsilon_a^\mu(k) \varepsilon_a^\nu(k)^* \right] \left[\sum_{a'=1,2} \varepsilon_{a'\mu}(k') \varepsilon_{a'\nu}(k')^* \right] =$$

$$\frac{2}{(k \cdot n)(k' \cdot n')} \left\{ (k \cdot n')(k' \cdot n) + (k \cdot k') \left[(n \cdot n') + \frac{(k' \cdot n)}{(k' \cdot n')} + \frac{(k \cdot n')}{(k \cdot n)} \right] + \frac{1}{2} \frac{(k \cdot k')^2}{(k \cdot n)(k' \cdot n')} \right\}.$$

Now let $n^\mu = (0, \hat{z})^T$, $k^\mu = \omega(1, \hat{z})^T$ and

$$k'^\mu = \omega' \begin{pmatrix} 1 \\ \hat{n}' \end{pmatrix}, \quad n'^\mu = \begin{pmatrix} 0 \\ \hat{n}' \end{pmatrix}, \quad \hat{n}' = \begin{pmatrix} \sin\theta \\ 0 \\ \cos\theta \end{pmatrix}.$$

We have:

$$k \cdot k' = \omega\omega'(1 - \cos\theta), \quad n \cdot n' = -\cos\theta$$

$$k \cdot n = -\omega, \quad k' \cdot n' = -\omega'$$

$$k \cdot n' = -\vec{k} \cdot \hat{n}' = -\omega\cos\theta, \quad k' \cdot n = -\vec{k}' \cdot \hat{n} = -\omega'\cos\theta.$$

Therefore (let $c \equiv \cos\theta$):

$$S = \frac{2}{\omega\omega'} \left\{ (\omega'c)(\omega c) + \omega\omega'(1-c) \left[-c + \frac{-\omega'c}{-\omega'} + \frac{-\omega c}{-\omega} \right] + \frac{1}{2} \frac{[\omega\omega'(1-c)]^2}{[-\omega][-\omega']} \right\}$$

$$= 2\left\{ c^2 + (1-c)c + \tfrac{1}{2}(1-c)^2 \right\}$$

$$= 2\{c^2 + c - c^2 + \tfrac{1}{2}(1 - 2c + c^2)\}$$

$$= 2\{c + \tfrac{1}{2} - c + \tfrac{1}{2}c^2\}$$

$$= 1 + c^2$$

$$= 2 - \sin^2\theta \ .$$

The problem wants us to take equation (9) on p. 154 and replace it with

$$\mathbb{M}^2 \equiv \tfrac{1}{2} \sum_{a=1,2} \sum_{a'=1,2} \left(\tfrac{1}{2} \sum\sum |\mathcal{M}|^2 \right)$$

$$= \frac{e^4}{(2m)^2} \left[2\left(\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - 2 \right) + 4 \times \tfrac{1}{2} S \right]$$

$$= \frac{e^4}{(2m)^2} \left[2\left(\frac{\omega'}{\omega} + \frac{\omega}{\omega'} \right) - 4 + 2(2 - \sin^2\theta) \right]$$

$$= \frac{e^4}{(2m)^2} \left[2\left(\frac{\omega'}{\omega} + \frac{\omega}{\omega'} \right) - 2\sin^2\theta \right]$$

$$= \frac{e^4}{(2m)^2} \left[2\left(\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\theta \right) \right] \ .$$

That is the result we were asked to prove.

2. Repeat the calculation of Compton scattering for circularly polarized photons.

Solution:

The explicit forms of the photon polarization vectors never entered into the derivation of the Klein-Nishina formula on p. 155; the only assumption was that the polarization vectors were real. For circularly polarized photons, everything goes through as before except that the ε_a^μ are complex. The result is equation (12) on p. 155 with the replacement $(\varepsilon\varepsilon')^2 \to |\varepsilon^*\varepsilon'|^2$.

III Renormalization and Gauge Invariance

III.1 Cutting Off Our Ignorance

1. Work through the manipulations leading to (9) without referring to the text.

Solution:

Start with equation (3):

$$\mathcal{M} = -i\lambda_{\text{bare}} + iC\lambda_{\text{bare}}^2 \left[\log\left(\frac{\Lambda^2}{s}\right) + \log\left(\frac{\Lambda^2}{t}\right) + \log\left(\frac{\Lambda^2}{u}\right) \right] + O(\lambda_{\text{bare}}^3)$$

Here we write the physical coupling as λ and the unphysical bare coupling as λ_{bare}.

We measure the physical coupling λ by performing some experiment at a particular center-of-mass energy and so on, and thus obtain equation (4):

$$-i\lambda = -i\lambda_{\text{bare}} + iC\lambda_{\text{bare}}^2 \left[\log\left(\frac{\Lambda^2}{s_0}\right) + \log\left(\frac{\Lambda^2}{t_0}\right) + \log\left(\frac{\Lambda^2}{u_0}\right) \right] + O(\lambda_{\text{bare}}^3)$$

We now pass to the notation of equations (5) and (6), where the sum of logs in the square brackets is denoted simply by L:

$$(5) \qquad \mathcal{M} = -i\lambda_{\text{bare}} + iC\lambda_{\text{bare}}^2 L + O(\lambda_{\text{bare}}^3)$$
$$(6) \qquad -i\lambda = -i\lambda_{\text{bare}} + iC\lambda_{\text{bare}}^2 L_0 + O(\lambda_{\text{bare}}^3)$$

Eliminate λ_{bare} by solving for it in terms of λ:

$$\begin{aligned}
-i\lambda_{\text{bare}} &= -i\lambda - iC\lambda_{\text{bare}}^2 L_0 + O(\lambda_{\text{bare}}^3) \\
&= -i\lambda + iC\left[-i\lambda + O(\lambda^2)\right]^2 L_0 + O(\lambda^3) \\
&= -i\lambda - iC\lambda^2 L_0 + O(\lambda^3)
\end{aligned}$$

Now plug this into (5) to get:

$$\begin{aligned}
\mathcal{M} &= \left(-i\lambda - iC\lambda^2 + O(\lambda^3)\right) - iC\left(-i\lambda + O(\lambda^2)\right)^2 L + O(\lambda^3) \\
&= -i\lambda - iC\lambda^2 L_0 + iC\lambda^2 L + O(\lambda^3) \\
&= -i\lambda + iC\lambda^2 (L - L_0) + O(\lambda^3) \\
&= -i\lambda + iC\lambda^2 \left[\log\left(\frac{s_0}{s}\right) + \log\left(\frac{t_0}{t}\right) + \log\left(\frac{u_0}{u}\right) \right]
\end{aligned}$$

Note that all dependence on the arbitrary cutoff has disappeared.

III.3 Counterterms and Physical Perturbation Theory

1. Show that in $(1+1)$-dimensional spacetime the Dirac field ψ has mass dimension $\frac{1}{2}$, and hence the Fermi coupling is dimensionless.

Solution:

The action $S = \int d^d x\, \mathcal{L}$ is dimensionless, so $\mathcal{L}$ has dimensions of $(\text{mass})^d$, denoted by $[\mathcal{L}] = d$. The derivative ∂_μ has mass dimension $+1$, so the kinetic term $\bar\psi \gamma^\mu \partial_\mu \psi$ gives:

$$\mathcal{L} = i\bar\psi\,\slashed\partial\psi \implies [\mathcal{L}] = [\partial] + 2[\psi] \implies d = 1 + 2[\psi] \implies [\psi] = \frac{d-1}{2}$$

Therefore, for $d = 1 + 1 = 2$, we have $[\psi] = 1/2$.

The Fermi interaction has the form $\mathcal{L} = G\,(\bar\psi\psi)(\bar\psi\psi)$. As above, this implies:

$$[\mathcal{L}] = [G] + 4[\psi] \implies d = [G] + 4\left(\frac{d-1}{2}\right) \implies [G] = 2 - d$$

In $d = 3 + 1 = 4$, $[G] = -2$, as discussed in the text. In $d = 1 + 1 = 2$, we have $[G] = 0$, so the Fermi coupling is dimensionless and thus the theory is renormalizable.

2. Derive (11) and (13).

Solution:

Consider first the Yukawa theory, defined by the Lagrangian density:

$$\mathcal{L} = \bar\psi(i\slashed\partial - m)\psi + \frac{1}{2}[(\partial\varphi)^2 - \mu^2\varphi^2]\underbrace{-\lambda\,\varphi^4 + f\,\varphi\bar\psi\psi}_{\equiv\;\mathcal{L}_{\text{int}}}$$

As in the book, we are really doing physical ("dressed") perturbation theory, in which case we should have renormalizing factors for each term or, equivalently, we should add a series of counterterms. For this problem, all that matters is the structure of the Lagrangian, so this issue is not important.

Introduce the sources J for φ, η for $\bar\psi$ and $\bar\eta$ for ψ and expand the path integral with sources in a Taylor series, as in equation (4) on page 44 and as described below equation (20)

on page 128:

$$\mathcal{Z}(\eta, \bar{\eta}, J) = \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}\varphi \, e^{iS[\psi, \bar{\psi}, \varphi] + i \int d^4 x (J\varphi + \bar{\eta}\psi + \bar{\psi}\eta)}$$

$$= e^{i \int d^4 x \, \mathcal{L}_{\text{int}}(\varphi \to -i\delta_J, \psi \to -i\delta_\eta, \bar{\psi} \to +i\delta_{\bar{\eta}})} \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}\varphi \, e^{iS[\psi, \bar{\psi}, \varphi]}$$

$$\sim \sum_{V_\lambda = 0}^{\infty} [\delta_J^4]^{V_\lambda} \sum_{V_f = 0}^{\infty} [\delta_J \delta_\eta \delta_{\bar{\eta}}]^{V_f} \sum_{B_I = 0}^{\infty} [J \, G_\varphi \, J]^{B_I} \sum_{F_I = 0}^{\infty} [\bar{\eta} \, G_\psi \, \eta]^{F_I}$$

The tilde means we are being schematic and ignoring all numerical factors, which are irrelevant for this problem. G_φ is the propagator for φ, and G_ψ is the propagator for ψ, both of which appear from performing the Gaussian integrals as usual. V_λ, V_f, B_I and B_F are, at this level, simply dummy summation variables.

These names are chosen for a reason: V_λ counts the number of $\lambda\varphi^4$ vertices, V_f counts the number of $f\varphi\bar{\psi}\psi$ vertices, B_I counts the number of scalar propagators ("internal boson lines"), and F_I counts the number of fermion propagators ("internal fermion lines").

Moreover, define $B_E \equiv \#$ Js killed, and $F_E \equiv \#$ ηs killed $+ \#$ $\bar{\eta}$s killed. How many Js are killed? Each δ_J kills one, and there are $4V_\lambda + V_f$ of them in a given term, as can be seen directly from the terms $[\delta_J^4]^{V_\lambda}$ and $[\delta_J...]^{V_f}$. How many Js persist? Each propagator term goes with two Js, so $2B_I$ persist, as can be seen from the term $[J...J]^{B_I}$. Thus $B_E = 4V_\lambda + V_f - 2B_I$, which matches equation (10) on page 163.

Similarly, each vertex kills V_f ηs, as can be read off from the term $[...\delta_\eta...]^{V_f}$, and each fermion propagator goes with one η, as can be read from $[...\eta]^{F_I}$, so the number of ηs killed is $V_f - F_I$. The number of $\bar{\eta}$s killed is the same, so $F_E = 2(V_f - F_I)$.

If the theory is renormalizable, then D should be independent of V_λ and V_f, so in anticipation of this let's solve for V_λ and V_f in terms of the other quantities:

$$V_f = F_I + \frac{1}{2}F_E$$

$$4V_\lambda = B_E + 2B_I - V_f = B_E + 2B_I - F_I - \frac{1}{2}F_E$$

Now we need to count powers of momentum. As explained on page 162, the number of loops L is the number of $\int d^4 k$ we have to perform, so the superficial degree of divergence D gets a $+4L$.

Each internal line carries a $\int d^4 k$, but each vertex carries a $\delta^4(k)$ that cancels those, so the number of $\int d^4 k$ that we actually have to perform is $L = B_I + F_I - (V_\lambda + V_f - 1)$, where the -1 accounts for the overall momentum conserving delta function. Each boson propagator goes like $1/k^2$, and each fermion propagator goes like $1/k$, so D gets a contribution of $-2B_I - F_I$. All together, we have:

$$D = 4L - 2B_I - F_I = 4[B_I + F_I - (V_\lambda + V_f - 1)] - 2B_I - F_I = 2B_I + 3F_I - 4(V_\lambda + V_f) + 4$$

76

Plugging in the values for V_f and V_λ gives:

$$D = 2B_I + 3F_I - \left(B_E + 2B_I - F_I - \frac{1}{2}F_E\right) - 4\left(F_I + \frac{1}{2}F_E\right) + 4$$

$$= (2-2)B_I + (4-4)F_I - B_E + \frac{1}{2}F_E - 4F_I - 2F_E + 4$$

$$= -B_E - \frac{3}{2}F_E + 4$$

Now consider the Fermi theory in (1+1) dimensions:

$$\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi - G\,(\bar{\psi}\psi)^2$$

The path integral with sources is

$$\mathcal{Z} \sim \sum_{V=0}^{\infty} [(\delta_\eta\delta_{\bar\eta})^2]^V \sum_{F_I=0}^{\infty} [\bar\eta\,G_\psi\eta]^{F_I}$$

Same procedure as above: $F_E \equiv 2\times(\#\ \eta\text{s killed}) = 2(2V - F_I) \implies 2V = \frac{1}{2}F_E + F_I$. This time, each loop is $\int d^2k$, so $D = 2L - F_I$, as stated above equation (13) on page 165. The number of loops is $L = F_I - (V-1) = F_I - \frac{1}{2}(\frac{1}{2}F_E + F_I) + 1 = \frac{1}{2}F_I - \frac{1}{4}F_E + 1$. So the superficial degree of divergence is:

$$D = 2\left(\frac{1}{2}F_I - \frac{1}{4}F_E + 1\right) - F_I = -\frac{1}{2}F_E + 2$$

That is equation (13) on page 165.

5. Show that the result $P = L - 1$ holds for all the theories we have studied.

Solution:

Take the QED Lagrangian rescaled by $\mathcal{L} \to \mathcal{L}/\hbar$

$$\mathcal{L} = \bar{\psi}[\frac{1}{\hbar}(i\partial\!\!\!/ - m)]\psi + \frac{1}{\hbar}eA_\mu\bar{\psi}\gamma^\mu\psi - \frac{1}{4\hbar}F_{\mu\nu}F^{\mu\nu}$$

Each fermion propagator contributes $\hbar$, each photon propagator contributes $\hbar$, and each fermion-photon vertex contributes $1/\hbar$. Therefore the total power of $\hbar$ in a diagram is $P = F_I + A_I - V$, where F_I is the number of internal fermion lines, A_I is the number of internal photon lines, and V is the number of vertices.

But the number of loops L is the number of $\int d^4k/[(2\pi)^4]$ we have to do, which is equal to the number of momenta flowing through each internal line minus the number of vertices, each of which conserves momentum with a $\delta^4(\sum k)$. So $L = F_I + A_I - V$, or in other words $P = L - 1$ again. It is clear that this happens for any theory.

III.5 Field Theory without Relativity

1. Obtain the Klein-Gordon equation for a particle in an electrostatic potential (such as that of the nucleus) by the gauge principle of replacing $(\partial/\partial t)$ in (2) by $\partial/\partial t - ieA_0$. Show that in the nonrelativistic limit this reduces to the Schrödinger's equation for a particle in an external potential.

Solution:

To add a photon to the relativistic free field theory $\mathcal{L} = (\partial\Phi^\dagger)(\partial\Phi) - m^2\Phi^\dagger\Phi$, replace the partial derivatives with covariant derivatives as $\partial_\mu\Phi \to D_\mu\Phi \equiv (\partial_\mu - ieA_\mu)\Phi$. (We also need the photon kinetic term $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, but it will play no role in what follows, so we drop it.) The Lagrangian is

$$\mathcal{L} = (D_\mu\Phi)^\dagger D^\mu\Phi + m^2\Phi^\dagger\Phi = \partial_\mu\Phi^\dagger\partial^\mu\Phi + m^2\Phi^\dagger\Phi + ieA^\mu(\Phi^\dagger\partial_\mu\Phi - \Phi\partial_\mu\Phi^\dagger) + e^2\Phi^\dagger\Phi A_\mu A^\mu$$

with a purely electric potential independent of time: $A^\mu(\vec{x}, t) = V(\vec{x})\delta^\mu{}_0$. Now take the non-relativistic limit as in the text:

$$\Phi(\vec{x}, t) = \frac{1}{\sqrt{2m}} e^{-imt}\varphi(\vec{x}, t) \implies \partial_t\Phi = \frac{1}{\sqrt{2m}}(-im\,\varphi + \partial_t\varphi)e^{-imt}$$

Therefore

$$\Phi^\dagger\partial_t\Phi = \frac{1}{\sqrt{2m}}e^{+imt}\varphi^\dagger\frac{1}{\sqrt{2m}}(-im\,\varphi + \partial_t\varphi)e^{-imt} = \frac{1}{2m}(-im\,\varphi^\dagger\varphi + \varphi^\dagger\partial_t\varphi)$$

and so we have

$$\Phi^\dagger\partial_t\Phi - \Phi\partial_t\Phi^\dagger = \frac{1}{2m}(-im\,\varphi^\dagger\varphi + \varphi^\dagger\partial_t\varphi) - \frac{1}{2m}(+im\,\varphi^\dagger\varphi + \varphi\partial_t\varphi^\dagger)$$

$$= -i\,\varphi^\dagger\varphi + \frac{1}{2m}(\varphi^\dagger\partial_t\varphi - \varphi\partial_t\varphi^\dagger)\ .$$

The term linear in the potential is therefore

$$ieA^\mu(\Phi^\dagger\partial_\mu\Phi - \Phi\partial_\mu\Phi^\dagger) = ieV(\vec{x})\left[-i\,\varphi^\dagger\varphi + \frac{1}{2m}(\varphi^\dagger\partial_t\varphi - \varphi\partial_t\varphi^\dagger)\right]$$

$$= eV(\vec{x})\varphi^\dagger\varphi + \frac{ie}{2m}V(\vec{x})(\varphi^\dagger\partial_t\varphi - \varphi\partial_t\varphi^\dagger)$$

$$= eV(\vec{x})\varphi^\dagger\varphi + \frac{ie}{m}V(\vec{x})\varphi^\dagger\partial_t\varphi + \text{(total time derivative)}$$

To this we add the term quadratic in the potential,

$$e^2\Phi^\dagger\Phi A_\mu A^\mu = \frac{e^2}{2m}\varphi^\dagger\varphi V(\vec{x})^2$$

Adapting the previously obtained Lagrangian from the book (equation (6) on p. 191), we obtain (up to total derivatives)

$$\mathcal{L} = \varphi^\dagger\left[i\partial_t + \frac{1}{2m}\nabla^2 + \frac{ie}{m}V(\vec{x})\partial_t + eV(\vec{x}) + \frac{e^2}{2m}V(\vec{x})^2\right]\varphi$$

78

Set $\frac{\delta S}{\delta \varphi^\dagger} = 0$ to get the equation of motion:

$$\left\{ \left[1 + \frac{e}{m}V(\vec{x})\right] i\partial_t + \frac{1}{2m}\nabla^2 + \left[1 + \frac{e}{2m}V(\vec{x})\right] eV(\vec{x}) \right\} \varphi(\vec{x}, t) = 0$$

In the non-relativistic limit $m\, \partial_t \varphi \ll \varphi$, and neglecting the term quadratic in the potential, we obtain

$$\left[i\partial_t + \frac{1}{2m}\nabla^2 + eV(\vec{x}) \right] \varphi(\vec{x}, t) = 0$$

which is precisely the Schrödinger equation $(i\partial_t + H)\varphi = 0$ with $H = \frac{\nabla^2}{2m} + eV$ the correct non-relativistic Hamiltonian for a particle in the presence of an external potential eV.

3. Given a field theory we can compute the scattering amplitude of two particles in the non-relativistic limit. We then postulate an interaction potential $U(\vec{x})$ between the two particles and use nonrelativistic quantum mechanics to calculate the scattering amplitude, for example in Born approximation. Comparing the two scattering amplitudes we can determine $U(\vec{x})$. Derive the Yukawa and the Coulomb potentials this way. The application of this method to the $\lambda(\Phi^\dagger\Phi)^2$ interaction is slightly problematic since the delta function interaction is a bit singular, but it should be all right for determining whether the force is repulsive or attractive.

Solution:

Given a scattering amplitude $\mathcal{M}$, the non-relativistic potential between distinguishable particles of masses m_1 and m_2 is

$$U(\vec{x}) = -\frac{1}{4m_1 m_2} \int \frac{d^3q}{(2\pi)^3} \mathcal{M}(\vec{q})\, e^{i\vec{q}\cdot\vec{x}} \ .$$

The Yukawa potential is derived in the text. Here we consider the $U(1)$-invariant scalar Lagrangian

$$\mathcal{L} = \partial\Phi^*\partial\Phi - m^2\Phi^*\Phi - \frac{\lambda}{4}(\Phi^*\Phi)^2 \ .$$

We would like to compute the potential between a Φ meson and an anti-meson. The scattering amplitude is just the constant $i\mathcal{M} = -i\lambda$, so the potential is

$$U(\vec{x}) = -\frac{1}{4m^2} \int \frac{d^3q}{(2\pi)^3}(-\lambda)\, e^{i\vec{q}\cdot\vec{x}} = \frac{+\lambda}{4m^2}\delta^3(\vec{x}) \ .$$

For $\lambda > 0$, this is a repulsive contact interaction. If you would like to consider a less singular Lagrangian, then write

$$\mathcal{L} = \partial\Phi^*\partial\Phi - m^2\Phi^*\Phi + \frac{1}{2}\left[(\partial\chi)^2 - M^2\chi^2\right] + iM\sqrt{\frac{\lambda}{2}}\chi\Phi^*\Phi$$

where χ is a real scalar field. Then consider the potential due to exchange of the field χ, which will reduce to the result obtained previously in the limit of taking the χ momentum to zero.

III.6 The Magnetic Moment of the Electron

3. By Lorentz invariance the right hand side of (7) has to be a vector. The only possibilities are $\bar{u}\gamma^\mu u$, $(p+p')^\mu \bar{u}u$, and $(p-p')^\mu \bar{u}u$. The last term is ruled out because it would not be consistent with current conservation. Show that the form given in (7) is in fact the most general allowed.

$$\langle p', s'|J^\mu(0)|p, s\rangle = \bar{u}(p', s') \left[\gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m} F_2(q^2) \right] u(p, s) \ , \quad q \equiv p' - p \tag{7}$$

Solution:

First, for notational convenience, write $u(p, s) = u$ and $u(p', s') = u'$. Lorentz invariance dictates that if we have one free μ index on the left, we need one free μ index on the right. The only such vectors we have are γ^μ, q^μ and $(p+p')^\mu$. However, since Lorentz invariance also dictates that we have no free spinor indices on the right, any vector we write down is going to be sandwiched between $\bar{u}'$ and u. Therefore, we may use the Gordon decomposition rearranged as:

$$\bar{u}'(p + p')^\mu u = \bar{u}' \left(2m\, \gamma^\mu - i\sigma^{\mu\nu}q_\nu \right) u$$

So between the spinors, $(p+p')$ is a linear combination of γ and q. Therefore anytime we write a function of $(p+p')$ we may as well write it as a function of q. Equation (7) follows immediately.

4. In chapter II.6, when discussing electron-proton scattering, we ignored the strong interaction that the proton participates in. Argue that the effects of the strong interaction could be included phenomenologically by replacing the vertex $\bar{u}(P, S)\gamma^\mu u(p, s)$ in (II.6.1) by

$$\langle P, S|J^\mu(0)|p, s\rangle = \bar{u}(P, S) \left[\gamma^\mu F_1(q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2m} F_2(q^2) \right] u(p, s) \tag{17}$$

Careful measurements of electron-proton scattering, thus determining the two proton form factors $F_1(q^2)$ and $F_2(q^2)$, earned R. Hofstadter the 1961 Nobel Prize. While we could account for the general behavior of these two form factors, we are still unable to calculate them from first principles (in contrast to the corresponding form factors for the electron.) See chapters IV.2 and VII.3.

Solution:

This form follows from parity and gauge invariance. Consult the literature, for example L. N. Hand, D. G. Miller and R. Wilson, "Electric and magnetic form factors of the nucleon," Rev. Mod. Phys, Vol. 35, No. 2, Apr. 1963 and J. D. Bjorken and E. A. Paschos, "Inelastic Electron-Proton and γ-Proton Scattering and the Structure of the Nucleon," Phys. Rev. Vol. 185, No. 3, 25 Sep. 1969. For a modern development regarding the form factors for the proton, see V. Barger, C. W. Chiang, W. Y. Keung and D. Marfatia, "Proton Size Anomaly," Phys. Rev. Lett. 106:153001, 2011 (arXiv:1011.3519v2 [hep-ph]).

III.7 Polarizing the Vacuum and Renormalizing the Charge

1. Calculate $\Pi_{\mu\nu}(q)$ using dimensional regularization.

Solution:

For this problem, we will use the metric signature $g = (-,+,+,+)$. The scheme of dimensional regularization is to calculate everything in d dimensions, then write $d = 4 - \varepsilon$ and take $\varepsilon \to 0^+$. The divergent piece will show up as a term proportional to $1/\varepsilon$. The polarization tensor is:

$$i\Pi^{\mu\nu}(k) = -\int \frac{d^d p}{(2\pi)^d} \frac{\text{Tr}\left[(-\slashed{k}-\slashed{p}+m)\,\gamma^\mu\,(-\slashed{p}+m)\,\gamma^\nu\right]}{[(k+p)^2 + m^2]\,[p^2 + m^2]}$$

$$= -\int \frac{d^d p}{(2\pi)^d} \frac{\text{Tr}\left[(\slashed{k}+\slashed{p})\gamma^\mu\slashed{p}\gamma^\nu + m^2\gamma^\mu\gamma^\nu\right]}{[(k+p)^2 + m^2]\,[p^2 + m^2]}$$

The simplification followed from remembering that the trace of an odd number of gamma matrices is zero. In d dimensions, the gamma matrices have dimension $2^{d/2}$, so the trace identities become:

$$\text{Tr}[\gamma^\mu\gamma^\nu] = -2^{d/2}g^{\mu\nu}$$

$$\text{Tr}[(\slashed{k}+\slashed{p})\gamma^\nu\slashed{p}\gamma^\nu] = 2^{d/2}[(k+p)^\mu p^\nu + (k+p)^\nu p^\mu - g^{\mu\nu}(k+p)\cdot p]$$

So the polarization tensor is now:

$$i\Pi^{\mu\nu}(k) = -2^{d/2}\int \frac{d^d p}{(2\pi)^d} \frac{(k+p)^\mu p^\nu + (k+p)^\nu p^\mu - g^{\mu\nu}(k+p)\cdot p - m^2 g^{\mu\nu}}{[(k+p)^2 + m^2]\,[p^2 + m^2]}$$

Play with the denominator using Feynman's formula:

$$\frac{1}{AB} = \int_0^1 dx\, \frac{1}{[xA + (1-x)B]^2}$$

Identifying $A = (k+p)^2 + m^2$ and $B = p^2 + m^2$, the polarization tensor is:

$$i\Pi^{\mu\nu}(k) = -2^{d/2} \int \frac{d^d p}{(2\pi)^d} \int_0^1 dx \, \frac{(k+p)^\mu p^\nu + (k+p)^\nu p^\mu - g^{\mu\nu}(k+p)\cdot p - m^2 g^{\mu\nu}}{\{x\,[(k+p)^2 + m^2] + (1-x)(p^2 + m^2)\}^2}$$

$$= -2^{d/2} \int \frac{d^d p}{(2\pi)^d} \int_0^1 dx \, \frac{(k+p)^\mu p^\nu + (k+p)^\nu p^\mu - g^{\mu\nu}(k+p)\cdot p - m^2 g^{\mu\nu}}{[x(k^2 + p^2 + 2k\cdot p + m^2) + p^2 + m^2 - xp^2 - xm^2]^2}$$

$$= -2^{d/2} \int \frac{d^d p}{(2\pi)^d} \int_0^1 dx \, \frac{(k+p)^\mu p^\nu + (k+p)^\nu p^\mu - g^{\mu\nu}(k+p)\cdot p - m^2 g^{\mu\nu}}{[p^2 + 2x\,k\cdot p + xk^2 + m^2]^2}$$

$$= -2^{d/2} \int \frac{d^d p}{(2\pi)^d} \int_0^1 dx \, \frac{(k+p)^\mu p^\nu + (k+p)^\nu p^\mu - g^{\mu\nu}(k+p)\cdot p - m^2 g^{\mu\nu}}{[(p+xk)^2 + x(1-x)k^2 + m^2]^2}$$

Now interchange the order of integration so that the $\int d^d p$ comes before the $\int dx$, and shift the integration variable to $q \equiv p + xk \implies p = q - xk$. For notational convenience, also define $\mathcal{D} \equiv x(1-x)k^2 + m^2$. We get:

$$i\Pi^{\mu\nu}(k) =$$
$$-2^{d/2} \int_0^1 dx \int \frac{d^d q}{(2\pi)^d} \, \frac{[q + (1-x)k]^\mu (q - xk)^\nu + (\mu \leftrightarrow \nu) - g^{\mu\nu}[q + (1-x)k]\cdot(q - xk) - m^2 g^{\mu\nu}}{[q^2 + \mathcal{D}]^2}$$
$$= -2^{d/2} \int_0^1 dx \int \frac{d^d q}{(2\pi)^d} \, \frac{2q^\mu q^\nu - 2x(1-x)k^\mu k^\nu - g^{\mu\nu}[q^2 - x(1-x)k^2] - m^2 g^{\mu\nu}}{[q^2 + \mathcal{D}]^2}$$

We have dropped terms in the numerator that are linear in q, because now the denominator is a function of the Lorentz-invariant q^2, so that $\int d^d q \, q^\mu/(q^2 + \mathcal{D})^2 = 0$. On a related note:

$$\int d^d q \, q^\mu q^\nu f(q^2) = C \, g^{\mu\nu} \int d^d q \, q^2 f(q^2) \quad \text{by Lorentz invariance}$$

$$g_{\mu\nu} \left[\int d^d q \, q^\mu q^\nu f(q^2) = C \, g^{\mu\nu} \int d^d q \, q^2 f(q^2) \right]$$

$$\int d^d q \, q^2 f(q^2) = C \, \mathrm{Tr}(\mathbb{I}_d) \int d^d q \, q^2 f(q^2)$$

$$1 = C \, d$$

$$\implies \int d^d q \, q^\mu q^\nu f(q^2) = \frac{1}{d} g^{\mu\nu} \int d^d q \, q^2 f(q^2)$$

Therefore, in the numerator of our integral we can substitute $q^\mu q^\nu \to g^{\mu\nu} q^2/d$ to get:

$$i\Pi^{\mu\nu}(k) =$$

$$= -2^{d/2} \int_0^1 dx \int \frac{d^d q}{(2\pi)^d} \frac{\frac{2}{d}q^2\, g^{\mu\nu} - 2x(1-x)k^\mu k^\nu - g^{\mu\nu}[q^2 - x(1-x)k^2] - m^2 g^{\mu\nu}}{[q^2 + \mathcal{D}]^2}$$

$$= -2^{d/2} \int_0^1 dx \int \frac{d^d q}{(2\pi)^d} \frac{2(\frac{1}{d}-1)\, q^2\, g^{\mu\nu} - 2x(1-x)k^\mu k^\nu + g^{\mu\nu}\, x(1-x)k^2 - m^2\, g^{\mu\nu}}{[q^2 + \mathcal{D}]^2}$$

$$= -2^{d/2} \int_0^1 dx \left\{ \left(\frac{2}{d}-1\right) g^{\mu\nu}\, \mathcal{I}_1 + \left[-2x(1-x)k^\mu k^\nu + g^{\mu\nu}\, x(1-x)k^2 - m^2\, g^{\mu\nu}\right] \mathcal{I}_0 \right\}$$

$$\mathcal{I}_n \equiv \int \frac{d^d q}{(2\pi)^d} \frac{(q^2)^n}{[q^2 + \mathcal{D}]^2} = \frac{i}{(4\pi)^{d/2}} \frac{\Gamma(2 - n - \frac{d}{2})\Gamma(n + \frac{d}{2})}{\Gamma(\frac{d}{2})} \mathcal{D}^{-(2-n-\frac{d}{2})}$$

$$= \frac{i}{16\pi^2} \frac{\Gamma(n + 2 - \frac{\varepsilon}{2})}{\Gamma(2 - \frac{\varepsilon}{2})} \Gamma(-n + \frac{\varepsilon}{2}) \mathcal{D}^n \left(\frac{4\pi}{\mathcal{D}}\right)^{\varepsilon/2}$$

As you can see, we have started to plug in $d = 4 - \varepsilon$. Before continuing, note a subtlety. In $d = 4$ spacetime dimensions, the gauge field A_μ has mass dimension $+1$ and the Dirac spinor ψ has mass dimension $+3/2$, so the gauge coupling e is dimensionless. But we are no longer working in $d = 4$ but rather in $d = 4 - \varepsilon$, so we need to find out how e changes with dimension. Taking typical terms in a Lagrangian, we get:

$$\mathcal{L} = -(\partial A)^2 \implies [A] = \frac{d-2}{2}$$

$$\mathcal{L} = i\bar{\psi}\,\slashed{\partial}\psi \implies [\psi] = \frac{d-1}{2}$$

$$\mathcal{L} = e\, A_\mu\, \bar{\psi}\gamma^\mu\psi \implies [e] = 2 - \frac{d}{2} = \frac{\varepsilon}{2}$$

To make this mass dimension explicit, we rewrite the gauge coupling as $e \to e\tilde{\mu}^{\varepsilon/2}$, where the parameter $\tilde{\mu}$ has dimensions of mass, so that e remains dimensionless for all d. The important point here is that since Π appears with an e^2, this consideration sends $\Pi \to \tilde{\mu}^\varepsilon\, \Pi$. So in the expression for the integral $\mathcal{I}_n$, the $(4\pi/\mathcal{D})^{\varepsilon/2}$ term is changed to $(4\pi\tilde{\mu}^2/\mathcal{D})^{\varepsilon/2}$.

Here are a few more intermediate steps[4]:

---

[4]Keeping $\varepsilon \neq 0$ in $2^{d/2} = 4 \times [1 + O(\varepsilon)]$ does not yield any new poles and is therefore not strictly necessary.

$$2^{d/2} = 2^{2-\varepsilon/2} = 4 \times 2^{-\varepsilon/2} = 4 \times \left(1 - \frac{\varepsilon}{2}\ln 2 + O(\varepsilon^2)\right)$$

$$\frac{2}{d} - 1 = \frac{2}{4-\varepsilon} - 1 = \frac{1}{2}\left(1 + \frac{\varepsilon}{4} + O(\varepsilon^2)\right) - 1 = -\frac{1}{2} + \frac{\varepsilon}{8} + O(\varepsilon^2)$$

$$\frac{\Gamma(3 - \frac{\varepsilon}{2})}{\Gamma(2 - \frac{\varepsilon}{2})} = \left(2 - \frac{\varepsilon}{2}\right)\frac{\Gamma(2 - \frac{\varepsilon}{2})}{\Gamma(2 - \frac{\varepsilon}{2})} = 2 - \frac{\varepsilon}{2}$$

$$\Gamma\left(-1 + \frac{\varepsilon}{2}\right) = -\frac{2}{\varepsilon} + \gamma - 1 + O(\varepsilon)$$

$$\Gamma\left(\frac{\varepsilon}{2}\right) = +\frac{2}{\varepsilon} - \gamma + O(\varepsilon)$$

$$\left(\frac{4\pi\tilde{\mu}^2}{\mathcal{D}}\right)^{\varepsilon/2} = 1 + \frac{\varepsilon}{2}\ln\left(\frac{4\pi\tilde{\mu}^2}{\mathcal{D}}\right) + O(\varepsilon^2)$$

Let's put these into the expressions for $\mathcal{I}_1$ and $\mathcal{I}_0$:

$$\begin{aligned}
\mathcal{I}_1 &= \frac{i}{16\pi^2}\left(2 - \frac{\varepsilon}{2}\right)\Gamma\left(-1 + \frac{\varepsilon}{2}\right)\mathcal{D}\left(\frac{4\pi\tilde{\mu}^2}{\mathcal{D}}\right)^{\varepsilon/2} \\
&= \frac{i}{8\pi^2}\left(1 - \frac{\varepsilon}{4}\right)\left[\frac{-2}{\varepsilon} + \gamma - 1 + O(\varepsilon)\right]\mathcal{D}\left[1 + \frac{\varepsilon}{2}\ln\left(\frac{4\pi\tilde{\mu}^2}{\mathcal{D}}\right) + O(\varepsilon^2)\right] \\
&= \frac{i}{8\pi^2}\left[\frac{-2}{\varepsilon} + \gamma - 1 + \frac{1}{2} + O(\varepsilon)\right]\mathcal{D}\left[1 + \frac{\varepsilon}{2}\ln\left(\frac{4\pi\tilde{\mu}^2}{\mathcal{D}}\right) + O(\varepsilon^2)\right] \\
&= \frac{i}{8\pi^2}\mathcal{D}\left[\frac{-2}{\varepsilon} - \ln\left(\frac{4\pi\tilde{\mu}^2}{\mathcal{D}}\right) + \gamma - 1 + O(\varepsilon)\right] \\
&= \frac{i}{8\pi^2}\mathcal{D}\left[\frac{-2}{\varepsilon} - \ln\left(\frac{4\pi\tilde{\mu}^2}{\mathcal{D}}\right) + \ln e^\gamma - \frac{1}{2} + O(\varepsilon)\right] \\
&= \frac{-i}{8\pi^2}\mathcal{D}\left[\frac{2}{\varepsilon} + \ln\left(\frac{4\pi\tilde{\mu}^2}{e^\gamma\mathcal{D}}\right) + \frac{1}{2} + O(\varepsilon)\right]
\end{aligned}$$

$$\begin{aligned}
\mathcal{I}_0 &= \frac{i}{16\pi^2}\Gamma\left(\frac{\varepsilon}{2}\right)\left(\frac{4\pi\tilde{\mu}^2}{\mathcal{D}}\right)^{\varepsilon/2} \\
&= \frac{i}{16\pi^2}\left[\frac{2}{\varepsilon} - \gamma + O(\varepsilon)\right]\left[1 + \frac{\varepsilon}{2}\ln\left(\frac{4\pi\tilde{\mu}^2}{\mathcal{D}}\right) + O(\varepsilon^2)\right] \\
&= \frac{i}{16\pi^2}\left[\frac{2}{\varepsilon} + \ln\left(\frac{4\pi\tilde{\mu}^2}{\mathcal{D}}\right) - \gamma + O(\varepsilon)\right] \\
&= \frac{i}{16\pi^2}\left[\frac{2}{\varepsilon} + \ln\left(\frac{4\pi\tilde{\mu}^2}{e^\gamma\mathcal{D}}\right) + O(\varepsilon)\right]
\end{aligned}$$

Define $\mu^2 \equiv 4\pi\tilde{\mu}^2/e^\gamma$ for convenience. After all, $\tilde{\mu}$ was defined arbitrarily anyway, just as long as it has dimensions of mass. Notice that now everything has an overall factor of i, so divide by that factor and plug in all of the results so far into the expression for the polarization tensor:

$$\Pi^{\mu\nu}(k) =$$
$$- 2^{d/2} \int_0^1 dx \left\{ \left(\frac{2}{d} - 1\right) g^{\mu\nu} \mathcal{I}_1/i + \left[-2x(1-x)k^\mu k^\nu + g^{\mu\nu} x(1-x)k^2 - m^2 g^{\mu\nu}\right] \mathcal{I}_0/i \right\}$$
$$= -[4 - \frac{\varepsilon}{2}\ln 2 + O(\varepsilon^2)] \int_0^1 dx \left\{ \left(\frac{-1}{2} + \frac{\varepsilon}{8}\right) g^{\mu\nu} \left(\frac{-1}{8\pi}\right) \mathcal{D} \left[\frac{2}{\varepsilon} + \ln\left(\frac{\mu^2}{\mathcal{D}}\right) + \frac{1}{2} + O(\varepsilon)\right] \right\}$$
$$- [4 - \frac{\varepsilon}{2}\ln 2 + O(\varepsilon^2)] \times$$
$$\int_0^1 dx \left\{ \left[-2x(1-x)k^\mu k^\nu + g^{\mu\nu} x(1-x)k^2 - m^2 g^{\mu\nu}\right] \left(\frac{1}{16\pi^2}\right) \left[\frac{2}{\varepsilon} + \ln\left(\frac{\mu^2}{\mathcal{D}}\right) + O(\varepsilon)\right] \right\}$$
$$= +\frac{1}{2\pi^2} \left[1 - \frac{\varepsilon}{8}\ln 2 + O(\varepsilon^2)\right] \int_0^1 dx \, g^{\mu\nu} \mathcal{D} \left\{-\frac{1}{\varepsilon} - \frac{1}{2}\ln\left(\frac{\mu^2}{\mathcal{D}}\right) - \frac{1}{4} + \frac{1}{4} + O(\varepsilon)\right\}$$
$$- \frac{1}{4\pi^2} \left[1 - \frac{\varepsilon}{8}\ln 2 + O(\varepsilon^2)\right] \int_0^1 dx \left[-2x(1-x)k^\mu k^\nu + g^{\mu\nu} x(1-x)k^2 - m^2 g^{\mu\nu}\right] \frac{2}{\varepsilon}$$
$$- \frac{1}{4\pi^2} \left[1 - \frac{\varepsilon}{8}\ln 2 + O(\varepsilon^2)\right] \int_0^1 dx \left[-2x(1-x)k^\mu k^\nu + g^{\mu\nu} x(1-x)k^2 - m^2 g^{\mu\nu}\right] \ln\left(\frac{\mu^2}{\mathcal{D}}\right)$$
$$+ O(\varepsilon)$$

Now group the terms in orders of ε:

$$\Pi^{\mu\nu}_{\varepsilon^{-1}}(k) =$$
$$\frac{1}{2\pi^2} \int_0^1 dx \left\{-g^{\mu\nu}\mathcal{D} + 2x(1-x)k^\mu k^\nu - g^{\mu\nu} x(1-x)k^2 + m^2 g^{\mu\nu}\right\} \frac{1}{\varepsilon}$$
$$= \frac{1}{2\pi^2} \int_0^1 dx \left\{-g^{\mu\nu}[x(1-x)k^2 + m^2] + 2x(1-x)k^\mu k^\nu - g^{\mu\nu} x(1-x)k^2 + m^2 g^{\mu\nu}\right\} \frac{1}{\varepsilon}$$
$$= \frac{1}{\pi^2} \int_0^1 dx \, x(1-x) \left\{k^\mu k^\nu - g^{\mu\nu} k^2\right\} \frac{1}{\varepsilon}$$
$$= \frac{1}{6\pi^2} \left(k^\mu k^\nu - g^{\mu\nu} k^2\right) \frac{1}{\varepsilon}$$

The divergent piece is indeed transverse. What about the finite piece?

$$\Pi^{\mu\nu}_{\varepsilon^0}(k) =$$

$$\frac{1}{2\pi^2} \int_0^1 dx\, g^{\mu\nu}\mathcal{D} \left[\frac{-1}{2}\ln\left(\frac{\mu^2}{\mathcal{D}}\right)\right] + \frac{\ln 2}{16\pi^2}g^{\mu\nu} \int_0^1 dx\, \mathcal{D}$$

$$+ \frac{\ln 2}{16\pi^2}\int_0^1 dx\, \left[-2x(1-x)k^\mu k^\nu + g^{\mu\nu}x(1-x)k^2 - m^2 g^{\mu\nu}\right]$$

$$- \frac{1}{4\pi^2}\int_0^1 dx\, \left[-2x(1-x)k^\mu k^\nu + g^{\mu\nu}x(1-x)k^2 - m^2 g^{\mu\nu}\right]\ln\left(\frac{\mu^2}{\mathcal{D}}\right)$$

$$= \frac{1}{2\pi^2}\int_0^1 dx\left[-\tfrac{1}{2}g^{\mu\nu}\left(x(1-x)k^2 + m^2\right) + x(1-x)k^\mu k^\nu - \tfrac{1}{2}g^{\mu\nu}x(1-x)k^2 + \tfrac{1}{2}m^2 g^{\mu\nu}\right]\ln\left(\frac{\mu^2}{\mathcal{D}}\right)$$

$$+ \frac{\ln 2}{16\pi^2}\left[g^{\mu\nu}\left(\tfrac{1}{6}k^2 + m^2\right) - \tfrac{1}{3}k^\mu k^\nu + \tfrac{1}{6}g^{\mu\nu}k^2 - m^2 g^{\mu\nu}\right]$$

$$= \frac{1}{2\pi^2}\int_0^1 dx\, x(1-x)\left[k^\mu k^\nu - k^2 g^{\mu\nu}\right]\ln\left(\frac{\mu^2}{\mathcal{D}}\right)$$

$$+ \frac{\ln 2}{48\pi^2}\left[k^2 g^{\mu\nu} - k^\mu k^\nu\right]$$

$$= \frac{1}{2\pi^2}\left(k^\mu k^\nu - k^2 g^{\mu\nu}\right)\left[\int_0^1 dx\, x(1-x)\ln\left(\frac{\mu^2}{\mathcal{D}}\right) + \frac{\ln 2}{24}\right]$$

The finite piece is also transverse. We have shown that dimensional regularization preserves the transverse structure of the polarization tensor and therefore preserves gauge invariance.

We have shown that $\Pi^{\mu\nu}(k) = (k^\mu k^\nu - k^2 g^{\mu\nu})\Pi(k^2)$, where

$$\Pi(k^2) = \frac{1}{6\pi^2\varepsilon} + \frac{1}{2\pi^2}\int_0^1 dx\, x(1-x)\ln\left(\frac{\mu^2}{\mathcal{D}}\right) + \frac{\ln 2}{24} + A$$

with $\mathcal{D} \equiv x(1-x)k^2 + m^2$, and with A being the counterterm. That is exactly the form of equation (13) on page 187 from using Pauli-Villars regularization.

To proceed with any physical calculation, we need to choose a renormalization scheme. Choose the "on-shell" renormalization scheme, in which the renormalized photon propagator has a pole at $k^2 = 0$ with residue 1, just like the tree-level propagator. In other words, we demand $\Pi_{\text{renormalized}}(k^2 = 0) = 0$. This fixes $A = -\frac{1}{6\pi\varepsilon} - \frac{1}{2\pi^2}\ln(\mu^2/m^2)\int_0^1 dx\, x(1-x) - \ln 2/24$, and therefore:

$$\Pi_{\text{renormalized}}(k^2) = \frac{1}{2\pi^2}\int_0^1 dx\, x(1-x)\ln\left(\frac{m^2}{x(1-x)k^2 + m^2}\right)$$

2. Study the modified Coulomb's law as determined by the Fourier integral $\int d^3q \, e^{i\vec{q}\cdot\vec{x}}/\{\vec{q}^2[1+e^2\Pi(\vec{q}^2)]\}$.

Solution:

The modified Coulomb potential is $V = V_{\text{coul}} + \delta V$, where

$$V_{\text{coul}}(\vec{r}) = \frac{e^2}{2\pi^2} \int d^3q \, \frac{e^{i\vec{q}\cdot\vec{r}}}{|\vec{q}|^2} = \frac{e^2}{r} \quad \text{(as usual)}$$

$$\delta V(\vec{r}) = -\left(\frac{e^2}{2\pi^2}\right)^2 \int d^3q \, \frac{e^{i\vec{q}\cdot\vec{r}}}{|\vec{q}|^2} \int_0^1 dx \, x(1-x) \ln\left[\frac{m^2}{m^2 + x(1-x)|\vec{q}|^2 - i\varepsilon}\right]$$

We have used the on-shell renormalized result for $\Pi(k^2)$ from the previous problem, setting $k^0 = 0$ and $\vec{k} = \vec{q}$.

First notice that the log has a branch cut at $|\vec{q}|^2/4 + m^2 < 0$, which causes the log to have a nonzero imaginary piece, which means that with enough energy a virtual electron-positron pair can become real. This can happen if the energy is greater than $2m$, which is precisely the branch point of the log. This is enough to show that there is a characteristic length scale of $1/(2m)$ in the potential.

Now let us compute the integrals explicitly. First evaluate the angular part of the Fourier integral:

$$\int d^3q \, \frac{e^{i\vec{q}\cdot\vec{r}}}{|\vec{q}|^2} f(|\vec{q}|^2) = 2\pi \int_0^\infty dq \, q^2 \, \frac{1}{q^2} f(q^2) \int_{-1}^1 d\eta \, e^{iqr\eta}$$

$$= 2\pi \int_0^\infty dq \, f(q^2) \, \frac{e^{iqr} - e^{-iqr}}{iqr}$$

Therefore we have $\delta V = -\left(\frac{e^2}{2\pi^2}\right)^2 \int_0^1 dx \, x(1-x) \, I$, where

$$I \equiv \int d^3q \, \frac{e^{i\vec{q}\cdot\vec{r}}}{|\vec{q}|^2} \ln\left(\frac{m^2}{m^2 + x(1-x)|\vec{q}|^2 - i\varepsilon}\right) = -\frac{2\pi i}{r} \int_0^\infty du \, \frac{e^{+iu} - e^{-iu}}{u} \ln\left(\frac{1}{1 + \beta u^2 - i\varepsilon}\right)$$

with $\beta \equiv x(1-x)/(mr)^2$. Now we will evaluate the integral over u by the method of contour integration. Consider the integral

$$J \equiv \oint_C dz \frac{e^{iz}}{z} \ln(1 + \beta z^2)$$

over the complex variable z, with the contour $C = \Gamma_2 + C_3 + \gamma_1 + C_2 + \gamma_2 + C_4 + \Gamma_1 + C_1$ as shown in the picture:
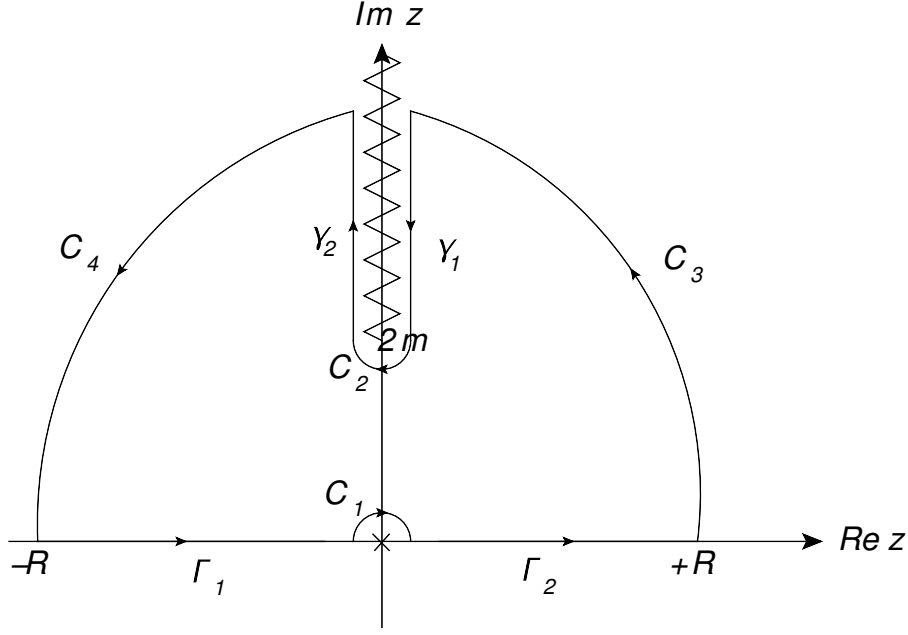
As stated previously and as shown in the diagram, the log has a branch cut starting at $z = i/\sqrt{\beta}$ extending along the imaginary axis to $z \to +i\infty$. Taking the radius R of the contour to infinity and the radius of the small semicircle C_1 to zero gives

$$\int_{\Gamma_1 + \Gamma_2} = \int_{-\infty}^{\infty} du \, \frac{e^{iu}}{u} \ln(1 + \beta u^2) \, .$$

The residue of the pole at C_1 is $\ln 1 = 0$, so $\int_{C_1} = 0$.

Along γ_2, the log function can be written $\ln z = \ln|z| + i\theta$, where $\theta = \pi/2$ is the location of the cut. The line γ_1 is on the other side of the cut, so the angle picks up an extra factor of 2π and hence on γ_1, we write $\ln z = \ln|z| + i(\theta + 2\pi)$.

Furthermore, along γ_2 the complex variable of integration is $z = 0^- + iy$, where y goes from R to $1/\sqrt{\beta}$. The zero indicates that we take γ_2 arbitrarily close to the left-hand side of the cut. Along γ_1, we have $z = 0^+ + iy$, where now y goes from $1/\sqrt{\beta}$ to R. Therefore, the only contribution to the integral from $\gamma_1 + \gamma_2$ is the part along the cut, since the rest cancels out. We have

$$\int_{\gamma_1 + \gamma_2} = -2\pi i \int_{1/\sqrt{\beta}}^{\infty} dy \, \frac{e^{-y}}{y}$$

where we have again taken $R \to \infty$. Since the contour does not enclose any residues, the whole contour integral J equals zero. Therefore:

$$\int_{-\infty}^{\infty} du \, \frac{e^{iu}}{u} \ln(1 + \beta u^2) = 2\pi i \int_{1/\sqrt{\beta}}^{\infty} dy \, \frac{e^{-y}}{y} \, .$$

We also need

$$\int_{-\infty}^{\infty} du \, \frac{e^{-iu}}{u} \ln(1 + \beta u^2) = \int_{+\infty}^{-\infty} (-du) \frac{e^{+iu}}{(-u)} \ln(1 + \beta u^2) = -\int_{-\infty}^{\infty} du \, \frac{e^{iu}}{u} \ln(1 + \beta u^2)$$

and
$$\int_0^\infty du \frac{e^{+iu} - e^{-iu}}{u} f(u^2) = \frac{1}{2} \int_{-\infty}^\infty du \frac{e^{+iu} - e^{-iu}}{u} f(u^2)$$

so therefore:
$$\int_{-\infty}^\infty du \frac{e^{iu} - e^{-iu}}{u} \ln(1 + \beta u^2) = 2\pi i \int_{1/\sqrt{\beta}}^\infty dy \frac{e^{-y}}{y} .$$

Thus we find
$$I = \frac{+2\pi i}{r} \left(2\pi i \int_{1/\sqrt{\beta}}^\infty dy \frac{e^{-y}}{y} \right) = -\frac{4\pi^2}{r} \int_{1/\sqrt{\beta}}^\infty dy \frac{e^{-y}}{y} .$$

The 1-loop correction to the Coulomb potential is now written as two nested integrals over the real parameters x and y:

$$\delta V(r) = +\frac{e^4}{\pi^2 r} \int_0^1 dx \, x(1-x) \int_{a(x)}^\infty dy \frac{e^{-y}}{y}$$

where $a(x) \equiv 1/\sqrt{\beta} = mr/\sqrt{x(1-x)}$. In the short-distance limit, we can Taylor expand around $mr = 0$ to get

$$\int_{a(x)}^\infty dy \frac{e^{-y}}{y} = -\ln(mr) + \frac{1}{2} \ln[x(1-x)] - \gamma + \frac{mr}{\sqrt{x(1-x)}} + O(mr)^2$$

where $\gamma \equiv -\int_0^\infty dx \, e^{-x} \ln x \approx 0.577$. Using the integrals

$$\int_0^1 dx \, x(1-x) \ln[x(1-x)] = -\frac{5}{18} , \quad \int_0^1 dx \sqrt{x(1-x)} = \frac{\pi}{8}$$

and $\int_0^1 dx \, x(1-x) = \frac{1}{6}$, we have:

$$\int_0^1 dx \, x(1-x) \int_{a(x)}^\infty dy \frac{e^{-y}}{y} = -\frac{1}{6} \left[\ln(mr) + \frac{5}{6} + \gamma - \frac{3\pi}{4} mr + O(mr)^2 \right]$$

Therefore, the correction to the Coulomb potential in the limit $mr \ll 1$ is

$$\delta V(r) \approx -\frac{e^4}{6\pi^2 r} \left[\ln(mr) + \frac{5}{6} + \gamma - \frac{3\pi}{4} mr \right] .$$

It is customary to express loop corrections in terms of $\alpha \equiv e^2/(4\pi)$, so we can rewrite the potential as

$$V(r) \approx \frac{e^2}{r} \left\{ 1 - \frac{2\alpha}{3\pi} \left[\ln(mr) + \frac{5}{6} + \gamma - \frac{3\pi}{4} mr \right] \right\} .$$

III.8   Becoming Imaginary and Conserving Probability

1. Evaluate the imaginary part of the vacuum polarization function, and by explicit calculation verify that it is related to the decay rate of a vector particle into an electron and a positron.

Solution:

We first evaluate the imaginary part of the vacuum polarization function by the method leading to equation (8) on p. 211. The vacuum polarization function is

$$\Pi(z) = -\frac{e^2}{2\pi^2}M^2 \int_0^1 dx\, x(1-x) \ln\left[1 - x(1-x)\frac{z}{m^2}\right]$$

where we have included a factor of e^2 from using the Lagrangian $\mathcal{L} = \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ instead of $\mathcal{L} = \frac{1}{4e^2}F_{\mu\nu}F^{\mu\nu}$, and we have included a factor of M^2 since we are shifting the pole from $k^2 = M^2$ rather than from $k^2 = 0$.

The log goes imaginary when its argument is negative. Since $\ln(-1) = i\pi$, we find

$$\text{Im}\Pi(\sigma + i\varepsilon) = -\frac{e^2}{2\pi^2}M^2 \int_0^1 dx\, x(1-x)(-\pi)\,\theta[x(1-x)\frac{\sigma}{m^2} - 1]$$

$$= +\frac{e^2}{2\pi}M^2 \int_{x_-}^{x_+} dx\, x(1-x)$$

$$= \frac{e^2}{4\pi}M^2\left(x^2 - \tfrac{2}{3}x^3\right)\Big|_{x_-}^{x_+} = \frac{1}{4\pi}M^2\left[x_+^2 - x_-^2 - \tfrac{2}{3}\left(x_+^3 - x_-^3\right)\right]$$

where $x_\pm \equiv \frac{1}{2}\left(1 \pm \sqrt{1 - 4m^2/\sigma}\right)$ are the two roots of the quadratic equation $x^2 - x + m^2/\sigma = 0$ found by setting the argument of the step function to zero. We have $x_+^2 - x_-^2 = \sqrt{1 - 4m^2/\sigma}$ and $x_+^3 - x_-^3 = \sqrt{1 - 4m^2/\sigma}\,(1 - m^2/\sigma)$, and so

$$x_+^2 - x_-^2 - \tfrac{2}{3}(x_+^3 - x_-^3) = \frac{1}{3}\sqrt{1 - \frac{4m^2}{\sigma}}\left(1 + \frac{2m^2}{\sigma}\right).$$

The decay rate is given on p. 212 as $\Gamma = \text{Im}\Pi(M^2)/M$, so we find

$$\Gamma = \frac{e^2}{12\pi}M\sqrt{1 - \frac{4m^2}{M^2}}\left(1 + \frac{2m^2}{M^2}\right).$$

Now we will calculate the decay rate of a vector particle into a fermion-antifermion pair using Feynman diagrams. Consider the Lagrangian for quantum electrodynamics with a massive photon:

$$\mathcal{L} = \bar{\psi}\,i\,\slashed{D}\psi - m\bar{\psi}\psi + \tfrac{1}{2}M^2 A_\mu A^\mu - \tfrac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

The Dirac field ψ represents the electron with mass m and electric charge $-e$, as specified by the covariant derivative $D_\mu\psi = \partial_\mu\psi - ie(-1)A_\mu\psi = \partial_\mu\psi + ieA_\mu\psi$. This implies a vertex $-ie\gamma^\mu$ and an amplitude $i\mathcal{M}(\gamma \to e^+e^-) = \epsilon_\mu^*(p_0)\bar{u}_2[-ie\gamma^\mu]v_1$. (Notation:

$0 = $ photon, $1 = $ outgoing positron, $2 = $ outgoing electron). The magnitude squared is
$$|\mathcal{M}|^2 = e^2\, \epsilon_\mu^*(p_0)\epsilon_\nu(p_0)(\bar{u}_2\gamma^\mu v_1)(\bar{v}_1\gamma^\nu u_2) = e^2\, \epsilon_\mu^*(p_0)\epsilon_\nu(p_0)\text{tr}[\gamma^\mu(v_1\bar{v}_1)\gamma^\nu(u_2\bar{u}_2)].$$

Next we sum over the spins of the outgoing fermions using $\sum_s u\bar{u} = (\not{p}+m)/(2m)$ and $\sum_s v\bar{v} = (\not{p}-m)/(2m)$, and average over the three spin states of the incoming massive photon using
$$\sum_a \epsilon^*_{(a)\mu}(p)\epsilon_{(a)\nu}(p) = -\eta_{\mu\nu} + \frac{p_\mu p_\nu}{M^2} .$$

The spin-summed amplitude squared $\mathbb{M}^2 \equiv \frac{1}{3}\sum_{a,s_1,s_2}|\mathcal{M}|^2$ is

$$\mathbb{M}^2 = -e^2\frac{1}{3}\left(\frac{1}{2m}\right)^2\left[\eta_{\mu\nu} - \frac{p_{0\mu}p_{0\nu}}{M^2}\right]\text{tr}\left[\gamma^\mu(\not{p}_1-m)\gamma^\nu(\not{p}_2+m)\right]$$
$$= -\frac{e^2}{12m^2}\left\{\text{tr}\left[\gamma^\mu(\not{p}_1-m)\gamma_\mu(\not{p}_2+m)\right] - \text{tr}\left[\gamma^0(\not{p}_1-m)\gamma^0(\not{p}_2+m)\right]\right\}$$

where we have used $p_0^\mu = (M,\vec{0})$ in the rest frame of the decaying photon.

Now we need some of the identities in Appendix D. Using $\{\gamma^\mu,\gamma^\nu\} = 2\eta^{\mu\nu} \implies \gamma^\mu\gamma_\mu = 4I$, $\text{tr}(\gamma^\mu\gamma^\nu) = 4\eta^{\mu\nu}$ and $\gamma^\mu\not{p}\gamma_\mu = -2\not{p}$, we have

$$\text{tr}[\gamma^\mu(\not{p}_1-m)\gamma_\mu(\not{p}_2+m)] = \text{tr}[(-2\not{p}_1-4m)(\not{p}_2+m)]$$
$$= \text{tr}[-2\not{p}_1\not{p}_2-4m^2]$$
$$= -8[p_1\cdot p_2+2m^2] .$$

Also using $(\gamma^0)^2 = I$ and
$$\text{tr}[\not{a}\not{b}\not{c}\not{d}] = 4[(a\cdot b)(c\cdot d) + (a\cdot d)(b\cdot c) - (a\cdot c)(b\cdot d)]$$

we obtain

$$\text{tr}[\gamma^0(\not{p}_1-m)\gamma^0(\not{p}_2+m)] = \text{tr}[\gamma^0\not{p}_1\gamma^0\not{p}_2-m^2 I]$$
$$= 4[2p_1^0 p_2^0 - p_1\cdot p_2 - m^2] .$$

Subtracting these two gives

$$\text{tr}\left[\gamma^\mu(\not{p}_1-m)\gamma_\mu(\not{p}_2+m)\right] - \text{tr}\left[\gamma^0(\not{p}_1-m)\gamma^0(\not{p}_2+m)\right]$$
$$= -8(p_1\cdot p_2+2m^2) - 4(2p_1^0 p_2^0 - p_1\cdot p_2 - m^2)$$
$$= -4[2(p_1\cdot p_2+2m^2) + 2p_1^0 p_2^0 - p_1\cdot p_2 - m^2]$$
$$= -4[p_1\cdot p_2 + 2p_1^0 p_2^0 + 3m^2]$$
$$= -4[3p_1^0 p_2^0 - \vec{p}_1\cdot\vec{p}_2 + 3m^2].$$

Therefore the squared amplitude is

$$\mathbb{M}^2 = \frac{e^2}{3m^2}\left(3p_1^0 p_2^0 - \vec{p}_1\cdot\vec{p}_2 + 3m^2\right)$$

where $p_i^0 = \sqrt{|\vec{p}_i|^2 + m^2}$.

The differential decay rate is

$$d\Gamma = \frac{1}{2M} \frac{d^3p_1}{(2\pi)^3 p_1^0/m} \frac{d^3p_2}{(2\pi)^3 p_2^0/m} (2\pi)^4 \delta^4(p_0 - p_1 - p_2) \mathbb{M}^2$$

where $\delta^4(p_0 - p_1 - p_2) = \delta(M - p_1^0 - p_2^0)\delta^3(\vec{p}_1 + \vec{p}_2)$ in the rest frame of the parent photon. Integrating over the 3-dimensional delta function sets $\vec{p}_2 = -\vec{p}_1$, for which the squared amplitude becomes

$$\mathbb{M}^2 = \frac{e^2}{3m^2}[3(p_1^0)^2 + \vec{p}_1^2 + 3m^2]$$

$$= \frac{e^2}{3m^2}[3(\vec{p}_1^2 + m^2) + \vec{p}_1^2 + 3m^2]$$

$$= \frac{e^2}{3m^2}[4\vec{p}_1^2 + 6m^2] .$$

The decay rate is

$$\Gamma = \frac{1}{2M} \frac{m^2}{(2\pi)^2} \int \frac{d^3p_1}{\vec{p}_1^2 + m^2} \delta(M - 2\sqrt{\vec{p}_1^2 + m^2}) \mathbb{M}^2|_{\vec{p}_2 = -\vec{p}_1}$$

$$= \frac{1}{2M} \frac{1}{4\pi^2} \frac{e^2}{3} \int \frac{d^3p_1}{\vec{p}_1^2 + m^2} \delta(M - 2\sqrt{\vec{p}_1^2 + m^2})(4\vec{p}_1^2 + 6m^2)$$

$$= \frac{e^2}{3\pi M} \int dp \frac{p^2}{p^2 + m^2} \delta(M - 2\sqrt{p^2 + m^2})(2p^2 + 3m^2)$$

where $p \equiv |\vec{p}_1|$. The delta function is

$$\delta(M - 2\sqrt{p^2 + m^2}) = \frac{M}{4p_*}\delta(p - p_*)$$

where $p_* = \sqrt{(M/2)^2 - m^2}$. The decay rate is

$$\Gamma = \frac{e^2}{3\pi M}\left(\frac{p_*^2}{p_*^2 + m^2}\right)\frac{M}{4p_*}(2p_*^2 + 3m^2)$$

$$= \frac{e^2}{6\pi}\left(\frac{p_*}{p_*^2 + m^2}\right)\left(p_*^2 + \frac{3}{2}m^2\right)$$

$$= \frac{e^2}{6\pi}\left(\frac{4p_*}{M^2}\right)\left(\frac{M^2}{4} - m^2 + \frac{3}{2}m^2\right)$$

$$= \frac{e^2}{6\pi}p_*\left(1 + \frac{2m^2}{M^2}\right)$$

$$= \frac{e^2}{12\pi}M\sqrt{1 - \frac{4m^2}{M^2}}\left(1 + \frac{2m^2}{M^2}\right) .$$

This agrees with the result obtained previously.

2. Suppose we add a $g\varphi^3$ term to our scalar φ^4 theory. Show that to order g^4 there is a "box diagram" contributing to meson scattering $p_1 + p_2 \to p_3 + p_4$ with the amplitude

$$\mathcal{I} = g^4 \int \frac{d^4 k}{(2\pi)^4} \frac{1}{(k^2 - m^2)[(k+p_2)^2 - m^2][(k-p_1)^2 - m^2][(k+p_2-p_3)^2 - m^2]}$$

(where m^2 above is understood as $m^2 - i\varepsilon$. Note that this is a typo in the problem in the text.)

Calculate the integral explicitly as a funtion of $s = (p_1 + p_2)^2$ and $t = (p_3 - p_2)^2$. Study the analyticity property of $\mathcal{I}$ as a function of s for fixed t. Evaluate the discontinuity of $\mathcal{I}$ across the cut and verify Cutkosky's cutting rule. Check that the optical theorem works. What about the analyticity property of $\mathcal{I}$ as a function of t for fixed s? And as a function of $u = (p_3 - p_1)^2$?

Solution:

We add the term $\mathcal{L} = -\frac{1}{3!} g\varphi^3$ to the Lagrangian, which generates the cubic vertex $-ig$. The diagram in question is displayed below:
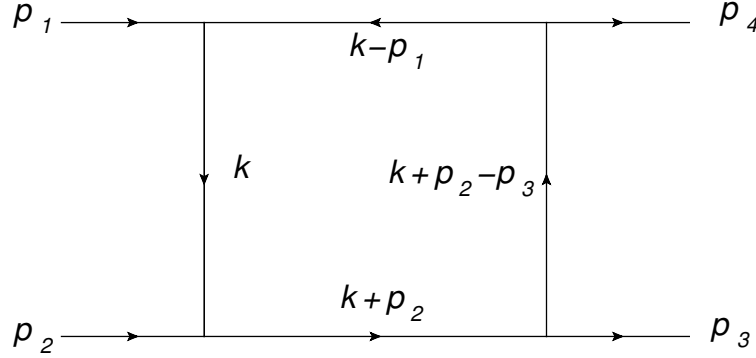
Each of the external momenta is on-shell, meaning $p_i^2 = m^2$ for $i = 1, 2, 3, 4$. Using the scalar propagator

$$i\Delta(p) = \frac{i}{p^2 - (m^2 - i\varepsilon)}$$

we obtain the amplitude

$$\mathcal{I} = (-ig)^4 \int \frac{d^4 k}{(2\pi)^4} [i\Delta(k)][i\Delta(k+p_2)][i\Delta(k-p_1)][i\Delta(k+p_2-p_3)]$$

$$= g^4 \int \frac{d^4 k}{(2\pi)^4} \frac{1}{[k^2 - m^2]} \frac{1}{[(k+p_2)^2 - m^2]} \frac{1}{[(k-p_1)^2 - m^2]} \frac{1}{[(k+p_2-p_3)^2 - m^2]}$$

where m^2 is understood as $m^2 - i\varepsilon$.

Note that the integral has four powers of k in the numerator and $4 \times 2 = 8$ powers of k in the denominator, so the integral goes as $\int dk\, k^3/[(k^2)^4] \sim \int dk/k^5$, which converges at high energy and thus does not need to be regularized.

Now that we have the integral, we should be clear as to what exactly this problem is asking. The steps are as follows:

First compute the integral using Feynman parameters, and evaluate the discontinuity of $\mathcal{I}$ across the cut at $s = 4m^2$. Second, "verify Cutkosky's cutting rule" by returning to the original integral, replacing the propagators with on-shell delta functions with positive frequency, evaluating the resulting integral and showing that we get the same result for the imaginary part as we got before.

Third, to "check the optical theorem," we are to compute the tree-level scattering amplitude obtained by cutting the box diagram down the middle, and plug it into the right-hand side of equation (19) on p. 215, which we repeat below for convenience:

$$\mathrm{Im}\mathcal{F}(i \to f) = \tfrac{1}{2} \sum_n (2\pi)^4 \delta^{(4)}(P_{ni}) \left(\prod_n \frac{1}{\rho_n^2} \right) \mathcal{F}(i \to n) \left[\mathcal{F}(f \to n) \right]^*$$

where in our case $\mathcal{F}(i \to f) = -i\mathcal{I}$. By evaluating the right-hand side explicitly, we are to observe that it equals the left-hand side, namely the imaginary part of the original box diagram that we computed previously in two different ways.

Finally, all results as a function of s for fixed t should be the same as those obtained as a function of t for fixed s.

To evaluate the integral directly, one introduces Feynman parameters and integrates over the loop momentum to put the integral in the form

$$\mathcal{I} = \frac{g^4}{16\pi^2} \int d^4x \, \frac{\delta(\sum_{i=1}^4 x_i - 1)\theta(x_1)\theta(x_2)\theta(x_3)\theta(x_4)}{[m^2(1 + x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1) + sx_1x_3 + tx_2x_4]} \, .$$

This is invariant under the interchange $s \leftrightarrow t$. Proceeding in this way, eventually one arrives at an expression in terms of logarithms and dilogarithms, with branch cuts in the appropriate kinematic channels ($s > 4m^2$ for fixed t, and $t > 4m^2$ for fixed s). For explicit details, see G. 't Hooft and M. Veltman, "Scalar One-Loop Integrals," Nucl. Phys. B153 (1979) 365-401 and A. Denner, U. Nierste and R. Scharf, "A Compact Expression for the Scalar One-Loop Four-Point Function," Nucl. Phys. B 367 (1991) 637-656.

We will now verify Cutkosky's cutting rule and the optical theorem.

We compute the imaginary part of the integral by replacing the cut internal propagators with momentum-conserving delta functions and thereby verify Cutkosky's cutting rule, as described on p. 215. For convenience, we collect the results from the example on pp. 216-217.

94

Given the amplitude

$$\Pi = +ig^2 \int \frac{d^4k}{(2\pi)^4} [i\Delta_{(\mu)}(k)][i\Delta_{(m)}(q-k)]$$

where the subscript on the scalar propagator denotes the location of the pole in momentum-squared (that is, the mass), Cutkosky tells us that twice the imaginary part of the amplitude is given by replacing the propagators $i\Delta$ with on-shell delta functions with positive frequency:

$$\text{Im}\,\Pi = \tfrac{1}{2}g^2 \int \frac{d^4k}{(2\pi)^4} \left[\theta(k^0)2\pi\delta(k^2-\mu^2)\right] \left[\theta\big((q-k)^0\big)\, 2\pi\delta\big((q-k)^2-m^2\big)\right]^2 .$$

As explained on p. 215, in our case we identify $\mathcal{I}$ with $i\Pi$. We first shift the loop momentum as $k \to k + p_1$ to obtain

$$\mathcal{I} = g^4 \int \frac{d^4k}{(2\pi)^4} \frac{1}{[k^2-m^2][(k+p_1)^2-m^2][(k+p_4)^2-m^2][(k+p_1+p_2)^2-m^2]}$$

where we have used $p_1 + p_2 - p_3 = p_4$. From this point on, we follow[5] P. van Nieuwenhuizen, "Muon-Electron Scattering Cross Section to Order α^3," Nucl. Phys. B28 (1971) 429-454.

We will cut the diagram through the propagators labeled by momenta k and $k + p_1 + p_2$. Cutkosky tells us that this entails replacing $i\Delta(k) \to \theta(k^0)2\pi\delta(k^2-m^2)$ and $i\Delta(k+p_1+p_2) \to \theta[-(k^0+p_1^0+p_2^0)]2\pi\delta[(k+p_1+p_2)^2-m^2]$. Note the minus sign in the second theta function; we are supposed to ensure that the momentum flows in one direction through the loop when we cut it. The imaginary part of the diagram $I \equiv \text{Im}(-i\mathcal{I})$ is

$$I = \frac{1}{2} \frac{g^4}{(2\pi)^2} \int d^4k \frac{\theta(k^0)\delta(k^2-m^2)\,\theta[-(k^0+p_1^0+p_2^0)]\delta[(k+p_1+p_2)^2-m^2]}{[(k+p_1)^2-m^2][(k+p_4)^2-m^2]} .$$

We can perform the integral over k^0 by using $\int dk^0\theta(k^0)\delta(k^2-m^2)f(k^0,\vec{k}) = \frac{1}{2\omega_k}f(\omega_k,\vec{k})$ where $\omega_k \equiv \sqrt{|\vec{k}|^2+m^2}$. Therefore:

$$I = \frac{g^4}{2(2\pi)^2} \int \frac{d^3k}{2\omega_k} \frac{\theta[-(k^0+p_1^0+p_2^0)]\delta[(p_1+p_2)^2+2k\cdot(p_1+p_2))]}{[p_1^2+2k\cdot p_1][p_4^2+2k\cdot p_4]} \bigg|_{k^0=\omega_k}$$

$$= \frac{g^4}{2(2\pi)^2} \int \frac{d^3k}{2\omega_k} \frac{\theta[-(\omega_k+\sqrt{s}\,)]\delta[s+2\omega_k\sqrt{s}\,]}{[m^2+\omega_k\sqrt{s}-2\vec{k}\cdot\vec{p_1}][m^2+\omega_k\sqrt{s}-2\vec{k}\cdot\vec{p_4}]}$$

where we have defined $s \equiv (p_1+p_2)^2$ and have chosen to work in the center-of-mass frame: $p_1^\mu = (\tfrac{1}{2}\sqrt{s},+\vec{p}),\ p_2^\mu = (\tfrac{1}{2}\sqrt{s},-\vec{p}) \implies \vec{p_1}+\vec{p_2}=0$.

---

[5] The reference uses the notation $q_m = p_1$, $q_e = p_2$, $q_e' = p_3$, $q_m = p_4$, $\Delta = k$ and $P = p_1 + p_2$ in Eq. (3.9). We treat the case for which all internal masses are equal, so $M = \lambda = m$. Also their metric is mostly positive, whereas ours is mostly negative.

Now let us define coordinates carefully in order to perform the integral. First fix the direction of the incoming momentum p_1 to point along the $\hat{z}$-axis:

$$\vec{p}_1 = |\vec{p}_1|\hat{z} \ , \quad |\vec{p}_1| = \tfrac{1}{2}\sqrt{s - 4m^2} \ .$$

Then fix the scattering plane as the (x, z)-plane, and define the scattering angle θ_0 in the center-of-mass frame as follows:

$$\vec{p}_4 = |\vec{p}_4| \left(\hat{x} \sin\theta_0 + \hat{z} \cos\theta_0\right) \ , \quad |\vec{p}_4| = |\vec{p}_1| = \tfrac{1}{2}\sqrt{s - 4m^2} \ .$$

We will perform the integral in spherical coordinates, so that the 3-vector to be integrated over is

$$\vec{k} = |\vec{k}| \left(\hat{x} \sin\theta \cos\varphi + \hat{y} \sin\theta \sin\varphi + \hat{z} \cos\theta\right) \ .$$

In these coordinates, we have

$$\vec{k} \cdot \vec{p}_1 = |\vec{k}||\vec{p}_1| \cos\theta$$

$$\vec{k} \cdot \vec{p}_4 = |\vec{k}||\vec{p}_4| \left(\sin\theta_0 \sin\theta \cos\varphi + \cos\theta_0 \cos\theta\right) \ .$$

Note that θ and φ are to be integrated over, while θ_0 is fixed.

Next use $\omega_k = \sqrt{k^2 + m^2}$ to rewrite the integral over $k \equiv |\vec{k}|$ as one over $\omega \equiv \omega_k$:

$$d^3k = d\Omega \, d\omega \, \omega \sqrt{\omega^2 - m^2}$$

where $d\Omega \equiv d\cos\theta \, d\varphi$. This will facilitate integration over the remaining delta function.

With these coordinates, the integral becomes:

$$I = \frac{g^4}{4(2\pi)^2} \int d\Omega \, d\omega \, \frac{\sqrt{\omega^2 - m^2} \, \theta[-(\omega + \sqrt{s})]\frac{1}{2\sqrt{s}}\delta(\tfrac{1}{2}\sqrt{s} + \omega)}{[\, m^2 + \omega\sqrt{s} - 2kp\cos\theta \,][\, m^2 + \omega\sqrt{s} - 2kp(\sin\theta_0 \sin\theta \cos\varphi + \cos\theta_0 \cos\theta)]}$$

where we have defined

$$k \equiv |\vec{k}| = \sqrt{\omega^2 - m^2} \quad \text{and} \quad p \equiv |\vec{p}_1| = |\vec{p}_4| = \tfrac{1}{2}\sqrt{s - 4m^2} \ .$$

The delta function sets $\omega = -\tfrac{1}{2}\sqrt{s}$, so that $k = \tfrac{1}{2}\sqrt{s - 4m^2} = p$, and therefore $kp = \tfrac{1}{4}(s - 4m^2)$. The integral is

$$I = \frac{g^4}{8(2\pi)^2} \sqrt{1 - \frac{4m^2}{s}}$$

$$\times \int d\Omega \frac{1}{[s - 2m^2 + (s - 4m^2)\cos\theta][s - 2m^2 + (s - 4m^2)(\sin\theta_0 \sin\theta \cos\varphi + \cos\theta_0 \cos\theta)]} \ .$$

The integral over angles is in the standard form[6]

$$\int d\Omega \frac{1}{(a + b\cos\theta)(A + B\cos\theta + C\sin\theta\cos\varphi)} = \frac{2\pi}{\sqrt{X}} \ln\left(\frac{aA - bB + \sqrt{X}}{aA - bB - \sqrt{X}}\right)$$

[6]See Appendix A of W. Beenakker and A. Denner, "Infrared Divergent Scalar Box Integrals with Applications in the Electroweak Standard Model," Nucl. Phys. B338 (1990) 349-370.

where

$$a = s - 2m^2$$
$$b = s - 4m^2$$
$$A = s - 2m^2 = a$$
$$B = (s - 4m^2)\cos\theta_0 = b\cos\theta_0$$
$$C = (s - 4m^2)\sin\theta_0 = b\sin\theta_0$$

and $X = (aA - bB)^2 - (a^2 - b^2)(A^2 - B^2 - C^2)$. Defining $t \equiv (p_1 - p_4)^2$ as usual implies $\cos\theta_0 = 1 + 2t/(s - 4m^2) = 1 + 2t/b$, so that for example $B = b + 2t$. The result can be brought into the form:

$$I = \frac{g^4}{4\pi} \frac{1}{P(s,t,m)} \ln\left(\frac{1 - Q(s,t,m)}{1 + Q(s,t,m)}\right)$$

where

$$P(s,t,m) \equiv t\sqrt{\left(1 - \frac{4m^2}{t}\right) s - 4m^2 \left(1 - \frac{3m^2}{t}\right)}$$

$$Q(s,t,m) \equiv \sqrt{1 - \frac{4m^2}{t}\left(\frac{s - 3m^2}{s - 4m^2}\right)} .$$

Finally, let us verify the optical theorem. If we cut the box diagram vertically, we arrive at the amplitudes $\mathcal{M}_L$ and $\mathcal{M}_R$ for the left and right halves respectively:
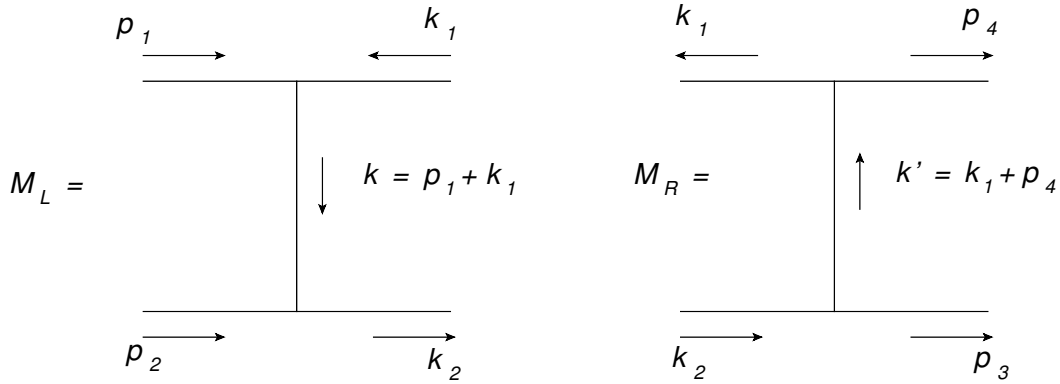
We have

$$\mathcal{M}_L = -i\frac{g^2}{(k_1 + p_1)^2 - m^2} , \quad \mathcal{M}_R = -i\frac{g^2}{(k_1 + p_4)^2 - m^2} .$$

We can also cut the diagram horizontally. After suitably shifting the internal momenta, this second cut will contribute an identical term to the imaginary part of the amplitude.

Define $\mathcal{F} \equiv -i\mathcal{M}$ as on p. 215. The sum $\sum_n$ over intermediate states becomes

$$\sum_n \to \int \frac{d^3k_1}{(2\pi)^3} \int \frac{d^3k_2}{(2\pi)^3}$$

and the product over normalization factors is

$$\prod_n \frac{1}{\rho_n^2} \to \frac{1}{(2\omega_1)(2\omega_2)}$$

where $\omega_i \equiv \sqrt{|\vec{k}_i|^2 + m^2}$. The optical theorem states that the imaginary part of the amplitude is given by

$$I = 2\times\frac{1}{2} \int \frac{d^3k_1}{(2\pi)^3} \int \frac{d^3k_2}{(2\pi)^3} \frac{1}{(2\omega_1)(2\omega_2)}(2\pi)^4\delta^4(k_1+k_2-p_1-p_2) \left[\frac{-g^2}{(k_1+p_1)^2-m^2}\right]\left[\frac{-g^2}{(k_1+p_4)^2-m^2}\right].$$

where the factor of accounts for the other possible cut, as discussed. In the center-of-mass frame, we have $\vec{p}_1 + \vec{p}_2 = 0$ and therefore

$$\delta^4(k_1 + k_2 - p_1 - p_2) = \delta(\omega_1 + \omega_2 - \sqrt{s})\,\delta^3(\vec{k}_1 + \vec{k}_2).$$

We can therefore do the integral over $\vec{k}_2$ and get

$$I = \frac{\pi g^4}{4(2\pi)^3} \int \frac{d^3k}{\omega_k^2} \delta(2\omega_k - \sqrt{s})\frac{1}{[(k+p_1)^2-m^2]}\frac{1}{[(k+p_4)^2-m^2]}$$

where $\omega_k \equiv \sqrt{|\vec{k}|^2 + m^2}$. As before, we can write $d^3k = d\Omega\, d\omega\, \omega\sqrt{\omega^2 - m^2}$ and integrate over the remaining delta function using $\delta(2\omega - \sqrt{s}) = \frac{1}{2}\delta(\omega - \frac{1}{2}\sqrt{s})$. We have:

$$I = \frac{g^4}{8(2\pi)^2}\frac{\sqrt{\omega^2 - m^2}}{\omega} \int d\Omega \frac{1}{[(k+p_1)^2-m^2]}\frac{1}{[(k+p_4)^2-m^2]}$$

$$= \frac{g^4}{8(2\pi)^2}\sqrt{1 - \frac{4m^2}{s}} \int d\Omega \frac{1}{[(k+p_1)^2-m^2]}\frac{1}{[(k+p_4)^2-m^2]}.$$

This is exactly the expression we got when applying the Cutkosky cutting rule.

3. Prove (28).
[Hint: Do unto $\int d^4x\, e^{iqx}\langle 0|[\mathcal{O}(x), \mathcal{O}(0)]|0\rangle$ what we did to $\int d^4x\, e^{iqx}\langle 0|T(\mathcal{O}(x)\mathcal{O}(0))|0\rangle$, namely insert $1 = \sum_n |n\rangle\langle n|$ (with $|n\rangle$ a complete set of states) between $\mathcal{O}(x)$ and $\mathcal{O}(0)$ in the commutator. Now we don't have to bother with representing the step function.]

Solution:

$$\int d^4x\, e^{iq\cdot x} \langle 0|[\mathcal{O}(x), \mathcal{O}(0)]|0\rangle = \int d^4x\, e^{iq\cdot x} \left(\langle 0|\mathcal{O}(x)\mathcal{O}(0)|0\rangle - \langle 0|\mathcal{O}(0)\mathcal{O}(x)|0\rangle\right)$$

$$= \int d^4x\, e^{iq\cdot x} \left(\langle 0|e^{+iP\cdot x}\mathcal{O}(0)e^{-iP\cdot x}\mathcal{O}(0)|0\rangle - \langle 0|\mathcal{O}(0)e^{+iP\cdot x}\mathcal{O}(0)e^{-iP\cdot x}|0\rangle\right)$$

$$= \int d^4x\, e^{iq\cdot x} \left(\langle 0|\mathcal{O}(0)e^{-iP\cdot x}\mathcal{O}(0)|0\rangle - \langle 0|\mathcal{O}(0)e^{+iP\cdot x}\mathcal{O}(0)|0\rangle\right)$$

$$= \int d^4x\, e^{iq\cdot x} \left(\langle 0|\mathcal{O}(0)\left(\sum_n |n\rangle\langle n|\right)e^{-iP\cdot x}\mathcal{O}(0)|0\rangle - \langle 0|\mathcal{O}(0)\left(\sum_n |n\rangle\langle n|\right)e^{+iP\cdot x}\mathcal{O}(0)|0\rangle\right)$$

$$= \int d^4x\, e^{iq\cdot x} \left(\sum_n \langle 0|\mathcal{O}(0)|n\rangle\langle n|\mathcal{O}(0)|0\rangle\, e^{-iP_n\cdot x} - \sum_n \langle 0|\mathcal{O}(0)|n\rangle\langle n|\mathcal{O}(0)|0\rangle\, e^{+iP_n\cdot x}\right)$$

$$= \sum_n |\langle 0|\mathcal{O}(0)|n\rangle|^2 \left(\int d^4x\, e^{i(q-P_n)\cdot x} - \int d^4x\, e^{i(q+P_n)\cdot x}\right)$$

$$= (2\pi)^4 \sum_n |\mathcal{O}_{0n}|^2 \left(\delta^{(4)}(q - P_n) - \delta^{(4)}(q + P_n)\right) \qquad (\dagger)$$

Equation (26) on p. 217 is

$$\mathrm{Im}\left(i\int d^4x\, e^{iq\cdot x}\langle 0|T(\mathcal{O}(x)\mathcal{O}(0))|0\rangle\right) = \frac{1}{2}(2\pi)^4 \sum_n |\mathcal{O}_{0n}|^2 \left(\delta^{(4)}(q - P_n) + \delta^{(4)}(q + P_n)\right)$$

We are told to prove equation (28) on p. 218, which is

$$\mathrm{Im}\left(i\int d^4x\, e^{iq\cdot x}\langle 0|T(\mathcal{O}(x)\mathcal{O}(0))|0\rangle\right) = \frac{1}{2}\int d^4x\, e^{iq\cdot x}\langle 0|[\mathcal{O}(x), \mathcal{O}(0)]|0\rangle \,.$$

We should now check equation $(\dagger)$ for each case on the right-hand side, namely for whether q^0 is positive or negative. For $q^0 > 0$, only the first delta function is nonzero and equation (28) follows immediately. For $q^0 < 0$, we have to track down a sign. For this we use translation invariance[7] in the form $\langle 0|[\mathcal{O}(x), \mathcal{O}(0)]|0\rangle = \langle 0|[\mathcal{O}(0), \mathcal{O}(-x)]|0\rangle$ to get

$$\int d^4x\, e^{iq\cdot x}\langle 0|[\mathcal{O}(x), \mathcal{O}(0)]|0\rangle = -\int d^4x\, e^{iq\cdot x}\langle 0|[\mathcal{O}(0), \mathcal{O}(x)]|0\rangle \qquad \text{(commutator)}$$

$$= -\int d^4x\, e^{-iq\cdot x}\langle 0|[\mathcal{O}(0), \mathcal{O}(-x)]|0\rangle \qquad (x \to -x)$$

$$= -\int d^4x\, e^{-iq\cdot x}\langle 0|[\mathcal{O}(x), \mathcal{O}(0)]|0\rangle \qquad \text{(translation invariance)} \,.$$

Thus for q satisfying the delta function $\delta^{(4)}(q+P_n)$, equation (28) also follows. This concludes the problem.

---

[7]Explicitly, use $\mathcal{O}(x) = e^{iP\cdot x}\mathcal{O}(0)e^{-iP\cdot x}$ and $P_\mu|0\rangle = 0$ to get $\langle 0|[\mathcal{O}(x), \mathcal{O}(0)]|0\rangle = \langle 0|\mathcal{O}(x)\mathcal{O}(0)|0\rangle - \langle 0|\mathcal{O}(0)\mathcal{O}(x)|0\rangle = \langle 0|\mathcal{O}(0)e^{-iP\cdot x}\mathcal{O}(0)|0\rangle - \langle 0|\mathcal{O}(0)e^{+iP\cdot x}\mathcal{O}(0)|0\rangle = \langle 0|\mathcal{O}(0)\mathcal{O}(-x)|0\rangle - \langle 0|\mathcal{O}(-x)\mathcal{O}(0)|0\rangle = \langle 0|[\mathcal{O}(0), \mathcal{O}(-x)]|0\rangle$.

IV Symmetry and Symmetry Breaking

IV.1 Symmetry Breaking

2. Construct the analog of (2) with N complex scalar fields and invariant under $SU(N)$. Count the number of Nambu-Goldstone bosons when one of the scalar fields acquires a vacuum expectation value.

$$\mathcal{L} = \frac{1}{2}\left[(\partial\vec{\varphi})^2 - m^2\vec{\varphi}^{\,2}\right] - \frac{1}{4}\lambda(\vec{\varphi}^{\,2})^2 \qquad (2)$$

Solution:

Let φ be a complex scalar field that transforms under the N-dimensional representation of $SU(N)$. Its Lagrangian is

$$\mathcal{L} = \partial_\mu\varphi_a^\dagger\partial^\mu\varphi^a + \mu^2\varphi_a^\dagger\varphi^a - \frac{\lambda}{2}(\varphi_a^\dagger\varphi^a)^2$$

The index a runs from 1 to N. Note that this Lagrangian has $U(N)$ symmetry, rather than just $SU(N)$ symmetry. Actually, this Lagrangian secretly has an even larger symmetry group that is at this level opaque because of our choice of field coordinates.

Decompose each field into its real and imaginary parts: $\varphi^a = \frac{1}{\sqrt{2}}(\chi^a + i\eta^a)$. The $U(N)$-invariant scalar product is

$$\varphi_a^\dagger\varphi^a = \frac{1}{2}(\chi_a - i\eta_a)(\chi^a + i\eta^a) = \frac{1}{2}(\chi_a\chi^a + \eta_a\eta^a)$$

Repackage the N χs and the N ηs into a $2N$-dimensional column vector:

$$\phi_A \equiv \begin{pmatrix} \vec{\chi} \\ \vec{\eta} \end{pmatrix}$$

Since the $U(N)$-invariant Lagrangian above depends only on this scalar product, it is actually invariant under the symmetry group $O(2N)$ (for more on this point, see p. 407):

$$\mathcal{L} = \sum_{A=1}^{2N}\frac{1}{2}\left[\partial_\mu\phi_A\partial^\mu\phi_A + \mu^2\phi_A\phi_A\right] - \frac{\lambda}{8}\left(\sum_{A=1}^{2N}\phi_A\phi_A\right)^2$$

The group $O(2N)$ has $2N(2N-1)/2 = N(2N-1)$ generators, as compared with the N^2 generators of $U(N)$. Now the problem is identical to IV.1.1, in which you are asked to show that $O(N)$ breaks to $O(N-1)$ and yields $N-1$ Nambu-Goldstone bosons. Before computing the Lagrangian explicitly, we first count the number of Nambu-Goldstone bosons from group theory, as described in the section "Counting Nambu-Goldstone bosons" on p. 199.

We start with the group $O(2N)$, which has $2N(2N-1)/2 = N(2N-1)$ generators. We break

it to $O(2N-1)$, which has $(2N-1)(2N-2)/2 = (2N-1)(N-1)$ generators. Therefore, the number of Nambu-Goldstone bosons is $N(2N-1) - (N-1)(2N-1) = 2N-1$.

Alternatively, if we break $U(N)$ to $U(N-1)$, we get $N^2 - (N-1)^2 = N^2 - (N^2 - 2N + 1) = 2N - 1$ Nambu-Goldstone bosons, which is the same number.

Now let us compute the Lagrangian explicitly. It is more straightforward to work with the explicitly $O(2N)$-invariant theory, repeated below

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi^A \partial^\mu \phi_A + \mu^2 \phi^A \phi_A) - \frac{\lambda}{8}(\phi^A \phi_A)^2$$

We have written indices up and down just for convenience; since $O(N)$ has the invariant tensor δ_{AB}, up and down make no difference. (Contrast with $SU(N)$, which does not have the invariant tensor δ_{AB} but only the invariant tensors $\varepsilon_{A_1...A_N}$, $\varepsilon^{A_1...A_N}$ and δ^A_B.)

Use the $O(2N)$ freedom to rotate the vacuum expectation value (VEV) of ϕ into the $2N^{\text{th}}$ component: $\langle \phi_A \rangle = v\delta_{A,2N}$. For clarity, let us define a new index $i = 1, ..., 2N - 1$ to single out the component with nonzero VEV. Let us also denote the shifted value of the last field by $h(x)$, so that we write:

$$\phi_{A=2N} = v + h(x) \ , \ \phi_{A \neq 2N} = \phi_i$$

The Lagrangian is:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu h \partial^\mu h + \partial_\mu \phi_i \partial^\mu \phi_i) + \frac{\mu^2}{2}\left[(v+h)^2 + \phi_i\phi_i\right] - \frac{\lambda}{8}\left[(v+h)^2 + \phi_i\phi_i\right]^2$$

We have:
$$(v+h)^2 = v^2 + 2vh + h^2$$

$$
\begin{aligned}
[(v+h)^2 + \phi_i\phi_i]^2 &= (v+h)^4 + 2(v+h)^2\phi_i\phi_i + (\phi_i\phi_i)^2 \\
&= (v^2 + 2vh + h^2)^2 + 2(v^2 + 2vh + h^2)\phi_i\phi_i + (\phi_i\phi_i)^2 \\
&= v^4 + 2v^2(2vh + h^2) + (2hv + h^2)^2 + 2v^2 + 4vh\phi_i\phi_i + 2h^2\phi_i\phi_i + (\phi_i\phi_i)^2 \\
&= v^4 + 4v^3h + 2v^2h^2 + 4v^2h^2 + 4vh^3 + h^4 + 2v^2\phi_i\phi_i + 4vh\phi_i\phi_i + 2h^2\phi_i\phi_i + (\phi_i\phi_i)^2 \\
&= v^4 + 4v^3h + 6v^2h^2 + 4vh^3 + h^4 + 2v^2\phi_i\phi_i + 4vh\phi_i\phi_i + 2h^2\phi_i\phi_i + (\phi_i\phi_i)^2
\end{aligned}
$$

We display this explicitly because the numerical factors are critical to get the right answer. The Lagrangian is:

$$\mathcal{L} = \frac{v^2}{2}\left(\mu^2 - \frac{\lambda v^2}{4}\right) + \frac{1}{2}(\partial_\mu h \partial^\mu h + \partial_\mu \phi_i \partial^\mu \phi_i)$$
$$+ \left[\frac{\mu^2}{2}2v - \frac{\lambda}{8}(4v^3)\right]h + \left[\frac{\mu^2}{2} - \frac{\lambda}{8}(6v^2)\right]h^2 + \left[\frac{\mu^2}{2} - \frac{\lambda}{8}(2v^2)\right]\phi_i\phi_i$$
$$- \frac{\lambda}{8}\left[4vh^3 + h^4 + 4vh\phi_i\phi_i + 2h^2\phi_i\phi_i + (\phi_i\phi_i)^2\right]$$

$$= \frac{v^2}{2}\left(\mu^2 - \frac{\lambda v^2}{4}\right) + \frac{1}{2}(\partial_\mu h \partial^\mu h + \partial_\mu \phi_i \partial^\mu \phi_i)$$
$$+ v\left[\mu^2 - \frac{\lambda v^2}{2}\right]h + \frac{1}{2}\left[\mu^2 - \frac{3\lambda v^2}{2}\right]h^2 + \frac{1}{2}\left[\mu^2 - \frac{\lambda v^2}{2}\right]\phi_i\phi_i$$
$$- \frac{\lambda v}{2}h^3 - \frac{\lambda v}{2}h\phi_i\phi_i - \frac{\lambda}{4}h^2\phi_i\phi_i - \frac{\lambda}{8}h^4 - \frac{\lambda}{8}(\phi_i\phi_i)^2$$

You see now what is going to happen: minimizing with respect to h at the point $h = 0$ and $\phi_i = 0$, along with $v \neq 0$, yields the condition $\mu^2 = \lambda v^2/2$. (Alternatively, you can minimize the Lagrangian with respect to v, set $h = \phi_i = 0$, then solve for v.) This also sets the coefficient of the term quadratic in ϕ_i to zero, resulting in $2N - 1$ massless Nambu-Goldstone bosons. As you can see, the resulting theory has $O(2N - 1)$ symmetry, along with cubic interaction terms, indicating that an h-particle can decay into two Nambu-Goldstone bosons.

IV.3   Effective Potential

2. Study V_{eff} in (1+1)-dimensional spacetime.

Solution:

The 1-loop effective potential for massive φ^4 theory in $d = 1 + 1$ dimensions is

$$V_{\text{eff}}(\varphi) = \frac{1}{2}(m^2 + B)\varphi^2 + \frac{1}{4!}\lambda\varphi^4 + \frac{i}{2}\int_\Lambda \frac{d^2k}{(2\pi)^2}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{\frac{1}{2}\lambda\varphi^2}{k^2 - m^2}\right)^n .$$

We will take $m^2 \to 0$ later. We have included a counterterm $\frac{1}{2}B\varphi^2$ to cancel a quadratic dependence on the cutoff Λ, but we have not included a counterterm $\frac{1}{4!}C\varphi^4$, because there is no quartic dependence on the cutoff. Only the $n = 1$ integral requires a finite UV cutoff:

$$\int_\Lambda \frac{d^2k}{(2\pi)^2}\frac{1}{k^2 - m^2} = -i\frac{1}{4\pi}\ln\left(\frac{\Lambda^2}{m^2}\right) .$$

In writing the right-hand side, we have dropped terms small compared to Λ, which we take arbitrarily large. For $n > 1$, we have

$$\int \frac{d^2k}{(2\pi)^2}\frac{1}{(k^2 - m^2)^n} = +i\frac{m^2}{4\pi}\frac{(-1)^n}{(n-1)(m^2)^n} .$$

Thus the regularized 1-loop effective potential is

$$V_{\text{eff}}(\varphi) = \frac{1}{2}(m^2 + B)\varphi^2 + \frac{1}{4!}\lambda\varphi^4 + \frac{1}{8\pi}\left[\frac{1}{2}\lambda\varphi^2 \ln\left(\frac{\Lambda^2}{m^2}\right) - m^2 S\right]$$

where we have defined the sum

$$S \equiv \sum_{n=2}^{\infty} \frac{(-1)^n}{n(n-1)}\left(\frac{\lambda\varphi^2}{2m^2}\right)^n = (1+x)\ln(1+x) - x , \quad x \equiv \frac{\lambda\varphi^2}{2m^2} .$$

We now require renormalization conditions. We would like to follow p. 241 in the text and impose "on-shell" renormalization conditions, for which $V_{\text{eff}}''(0) = m^2$. This fixes

$$B = -\frac{\lambda}{8\pi}\ln\left(\frac{\Lambda^2}{m^2}\right) .$$

Putting this B back into the effective potential, we find

$$V_{\text{eff}}(\varphi) = \frac{1}{2}m^2\varphi^2 + \frac{1}{4!}\lambda\varphi^4 + \frac{1}{16\pi}\left[\lambda\varphi^2 - (\lambda\varphi^2 + 2m^2)\ln\left(\frac{\lambda\varphi^2}{2m^2} + 1\right)\right] .$$

Taking m^2 arbitrarily small but nonzero, we have

$$V_{\text{eff}}(\varphi) = \frac{1}{4!}\lambda\varphi^4 - \frac{1}{16\pi^2}\lambda\varphi^2 \ln\left(\frac{\lambda\varphi^2}{2m^2}\right) .$$

For $\varphi \to 0$, the potential behaves as $V_{\text{eff}}(\varphi) \to +\frac{1}{16\pi^2}\lambda\varphi^2 \ln\left(\frac{2m^2}{\lambda\varphi^2}\right) > 0$. The $\mathbb{Z}_2 : \varphi \to -\varphi$ symmetry of the classical potential is not spontaneously broken.[8]

Now consider the case for which we start with a massless theory, for which the 1-loop effective potential is

$$V_{\text{eff}}(\varphi) = \frac{1}{4!}\lambda\varphi^4 + \frac{1}{2}B\varphi^2 + \frac{i}{2}\int\frac{d^2k}{(2\pi)^2}\sum_{n=1}^{\infty}\frac{1}{n}\left(\frac{\lambda\varphi^2}{2k^2}\right)^n$$

$$= \frac{1}{4!}\lambda\varphi^4 + \frac{1}{2}B\varphi^2 + \frac{1}{2}\int\frac{d^2k_E}{(2\pi)^2}\ln\left(1 + \frac{\lambda\varphi^2}{2k_E^2}\right)$$

where in the second line we have summed the series and rotated to Euclidean momentum. Regularizing the integral with an arbitrarily large UV cutoff Λ, we have

$$\int\frac{d^2k_E}{(2\pi)^2}\ln\left(1 + \frac{\lambda\varphi^2}{2k_E^2}\right) = \frac{1}{8\pi}\lambda\varphi^2 \ln\left(\frac{2\Lambda^2}{\lambda\varphi^2}\right)$$

where we have dropped terms that go to zero for $\Lambda \to \infty$. The regularized effective potential is

$$V_{\text{eff}}(\varphi) = \frac{1}{4!}\lambda\varphi^4 + \frac{1}{2}B\varphi^2 + \frac{1}{16\pi}\lambda\varphi^2 \ln\left(\frac{2\Lambda^2}{\lambda\varphi^2}\right) .$$

---

[8]Actually the $\varphi \to -\varphi$ symmetry is not broken in $(3+1)$ dimensions either. See the addendum at the end of this section.

The second derivative of this has a log divergence at $\varphi = 0$. Instead, define the curvature at an arbitrary field point $\varphi = \mu$ as $V_{\text{eff}}''(\mu) \equiv m^2(\mu)$. This fixes

$$B = m^2(\mu) - \frac{\lambda}{8\pi}\left[\ln\left(\frac{2\Lambda^2}{\lambda\varphi^2}\right) + 4\pi\mu^2 - 3\right] .$$

Note that in $(1{+}1)$ dimensions, φ is dimensionless so μ is dimensionless, and λ has dimensions of mass-squared. Putting B back into the effective potential gives

$$V_{\text{eff}}(\varphi) = \frac{1}{4!}\lambda\varphi^4 + \frac{1}{2}\left[m^2(\mu) - \tfrac{1}{2}\lambda\mu^2\right]\varphi^2 + \frac{1}{16\pi}\lambda\varphi^2\left[\ln\left(\frac{\mu^2}{\varphi^2}\right) + 3\right] .$$

This is now independent of Λ, as it should be by renormalizability.[9] At this point, if we take $\varphi \to 0$ we find $V_{\text{eff}}(\varphi) \to +\frac{\lambda}{16\pi}\varphi^2\ln\left(\frac{\mu^2}{\varphi^2}\right) > 0$. However, in this limit the log becomes large, and since the expansion parameter is $\lambda\ln(\mu^2/\varphi^2)$, perturbation theory is not trustworthy.

To obtain an improved perturbation expansion, we use the renormalization group. The effective potential must be independent of the arbitrary point μ, so differentiating the equation for $V_{\text{eff}}(\varphi)$ with respect to μ and setting the result to zero gives a flow equation for the parameter $m^2(\mu)$:

$$\mu\frac{d}{d\mu}m^2(\mu) = +\left(\mu^2 - \frac{1}{4\pi}\right)\lambda .$$

In $(1+1)$ dimensions, λ does not run and is therefore a constant with respect to μ. Thus this equation can be readily integrated (from an arbitrary point μ_0):

$$m^2(\mu) = m^2(\mu_0) + \lambda\left[\frac{1}{2}(\mu^2 - \mu_0^2) - \frac{1}{4\pi}\ln\left(\frac{\mu}{\mu_0}\right)\right] .$$

Parametrize the RG integration constant as $\mu_0^2 = \xi\varphi^2$. Then

$$V_{\text{eff}}(\varphi) = \tfrac{1}{2}m_{\text{eff}}^2(\xi)\,\varphi^2 + \tfrac{1}{4!}\lambda_{\text{eff}}(\xi)\,\varphi^4 , \quad \text{where:}$$

- $m_{\text{eff}}^2(\xi) \equiv m^2(\mu = \xi^{1/2}\varphi) + \frac{\lambda}{8\pi}(3 + \ln\xi) ,$

- $\lambda_{\text{eff}}(\xi) \equiv (1 - 12\xi)\lambda .$

The idea is to approach the origin $\varphi = 0$ carefully by keeping ξ small but nonzero, taking $\varphi \to 0$ and then taking $\xi \to 0$.

---

[9]Note also that, as in $d = (3 + 1)$ dimensions, all logs of λ have disappeared after renormalization. In Appendix A.2 of the Coleman-Weinberg paper, it is shown that by rescaling the loop momenta $k \to \lambda^{1/2}k$, and therefore $d^dk \to \lambda^{d/2}d^dk$, the contribution to the effective potential is of the form $\varphi^4 f(\varphi/M)\lambda^{V+\frac{d}{2}L-I}$, where V is the number of vertices, L is the number of loops, and I is the number of internal lines. For $d = 4$, the contribution goes as $\lambda^{V+2L-I} = \lambda^{L+1}$, and thus the 1-loop correction goes as λ^2, as shown in the text. For the present case of $d = 2$, the contribution goes as $\lambda^{V+L-I} = \lambda$, which is independent of the number of loops L. This explains our result that the 1-loop correction to massless φ^4 theory in $d = (1+1)$ dimensions goes as λ.

We are interested in "massless" φ^4 theory, so we want $m^2(\mu \to 0) \to 0$. (i.e., $V''_{\text{eff}}(0) = 0$.) From the RG equation, we have

$$m^2(\varphi) = m^2(\xi^{1/2}\varphi) + \lambda[\tfrac{1}{2}(1-\xi)\varphi^2 + \frac{1}{8\pi}\ln\xi] \, .$$

Consider $\xi > 0$, but $\varphi \to 0$. Then $m^2(\varphi) \approx m^2(\xi^{1/2}\varphi) + \frac{\lambda}{8\pi}\ln\xi$, which implies $m^2_{\text{eff}}(\xi) \approx m^2(\varphi) + \frac{3}{8\pi}\lambda$. If $m^2(\varphi \to 0) \to 0$, then $m^2_{\text{eff}}(\xi) \to \frac{3}{8\pi}\lambda > 0$.

So near the origin $\varphi \to 0$, taking $0 < \xi \ll \frac{1}{12}$ in massless φ^4 theory implies that $V_{\text{eff}}(\varphi) \approx \frac{1}{4!}\lambda\varphi^4$. The $\mathbb{Z}_2$ symmetry is not spontaneously broken.

5. Consider the electrodynamics of a complex scalar field

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + [(\partial^\mu + ieA^\mu)\varphi^\dagger][(\partial_\mu - ieA_\mu)\varphi] + \mu^2\varphi^\dagger\varphi - \lambda(\varphi^\dagger\varphi)^2.$$

In a universe suffused with the scalar field $\varphi(x)$ taking on the value φ independent of x as in the text, the Lagrangian will contain a term $(e^2\varphi^\dagger\varphi)A_\mu A^\mu$ so that the effective mass squared of the photon field becomes $M(\varphi)^2 \equiv 2e^2\varphi^\dagger\varphi$. Show that its contribution to $V_{\text{eff}}(\varphi)$ has the form

$$\int \frac{d^4k}{(2\pi)^4} \ln\left(\frac{k^2 - M(\varphi)^2}{k^2}\right).$$

Compare with (14) and (26). [Hint: Use the Landau gauge to simplify the calculation.] If you need help, I strongly urge you to read S. Coleman and E. Weinberg, *Phys. Rev.* D7: 1883, 1973, a paragon of clarity in exposition.

Solution:

In the Landau gauge (see p. 267 with $\xi = 0$), computing the photon contribution to the effective potential constitutes summing up the one-loop diagrams depicted below:
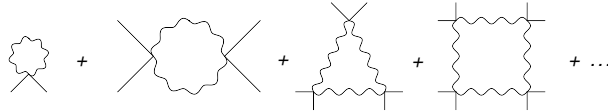
The external lines are scalar lines with no external momenta. The diagrams (from left to right) contain $n = 1, 2, 3, 4, \dots$ powers of the photon propagator. The reason for choosing Landau gauge is that otherwise there would be diagrams with both photons and scalars in the internal lines.

The massive photon propagator is (suppressing the $i\epsilon$):

$$i\Delta_{\mu\nu}(k) = \frac{-i}{k^2 - M(\varphi)^2}\left(\eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right)$$
$$= -i\Delta(k)P_{\mu\nu}(k)$$

105

where we have recognized the scalar propagator

$$i\Delta(k) = \frac{i}{k^2 - M(\varphi)^2}$$

and we have defined a projection operator

$$P_{\mu\nu}(k) \equiv \eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \ .$$

This is a projection operator in the sense that $P_{\mu\nu}(k)P^{\nu\rho}(k) = P_\mu{}^\rho(k)$, so that $\mathrm{tr}(P^n) = \mathrm{tr}P = 3$ for any positive integer n.

Because of this, and since every diagram in the above one-loop expansion contains an equal number of $e^2\varphi^\dagger\varphi A^\mu A_\mu \to 2ie^2\eta^{\mu\nu}$ vertices and photon propagators $i\Delta_{\mu\nu}(k) = -i\Delta(k)P_{\mu\nu}(k)$, we can simply replace the series by an equivalent series with internal scalar lines instead of photon lines, as long as we include a minus sign in the vertex and multiply the sum by an overall factor of 3. In terms of the scalar coupling λ, the replacement is $\lambda \to 2e^2$.

In other words, we can immediately jump to equation (14) in the text with the replacement $V''(\varphi) = M(\varphi)$ and an overall factor of 3 multiplying the integral.

Therefore, the photon results in an effective potential

$$V_{\mathrm{eff}}(\varphi) = V(\varphi) - 3 \times \tfrac{1}{2} i \int \frac{d^4k}{(2\pi)^4} \ln\left(\frac{k^2 - M(\varphi)^2}{k^2}\right) \ .$$

The photon contribution is equal to $(+3)$ real scalar contributions, which makes sense given the 3 polarization states of a massive vector boson (the photon obtains an effective mass in the φ background).

Now compare this to the contribution of a fermion (equation (26) on p. 243, with slight notational adjustments):

$$V_{\mathrm{eff}}(\varphi) = V(\varphi) + 4 \times \tfrac{1}{2} i \int \frac{d^4k}{(2\pi)^4} \ln\left(\frac{k^2 - M(\varphi)^2}{k^2}\right) \ .$$

The Dirac fermion contribution is equal to (-4) real scalar contributions to the effective potential.

Addendum 1: Effective Potential Revisited

On p. 242 of the text, it is shown that the one-loop correction $+\varphi^4 \ln \varphi^2$ overwhelms the classical $+\varphi^4$ potential near $\varphi = 0$. However, the conclusion to draw from this is not that quantum fluctuations break the discrete symmetry $\varphi \to -\varphi$ but rather that perturbation

106

theory is not valid near $\varphi = 0$. The issue is not whether $\hbar$ is small (indeed we have set it to 1), but rather, as explained on p. 242, that the expansion parameter is $\lambda \ln \varphi$ rather than just λ. For small φ, the second term in the expansion becomes larger than the first term, and the third term will be larger than second term, and so forth, which means that the perturbation expansion breaks down near $\varphi = 0$ and cannot tell us whether $\varphi = 0$ is a maximum or a minimum of the potential.

We now use the renormalization group to obtain a perturbative expansion that is valid near $\varphi = 0$ and show that the $\varphi \to -\varphi$ symmetry is not spontaneously broken by one-loop effects.[10]

The one-loop beta function for the quartic coupling λ is (equation (19) on p. 242)

$$M \frac{d\lambda(M)}{dM} = \frac{3}{16\pi^2} [\lambda(M)]^2 + O[\lambda(M)]^3 .$$

The solution to this equation, integrated from some arbitrary scale M_0, is

$$\lambda(M) = \frac{\lambda(M_0)}{1 - \frac{3}{32\pi^2} \lambda(M_0) \ln \left(\frac{M^2}{M_0^2} \right)} .$$

In the text we found the effective potential

$$V_{\text{eff}}(\varphi_c) = \frac{1}{4!} \lambda(M) \varphi_c^4 + \frac{1}{256\pi^2} [\lambda(M)]^2 \varphi_c^4 \left[\ln \left(\frac{\varphi_c^2}{M^2} \right) - \frac{25}{6} \right] .$$

The integration constant M_0 was arbitrary; let us choose the value $M_0 = \varphi_c$. Then the solution of the renormalization group equation for λ gives

$$\lambda(M) = \lambda(\varphi_c) \left\{ 1 + \frac{3}{32\pi^2} \lambda(\varphi_c) \ln \left(\frac{M^2}{\varphi_c^2} \right) + O \left[\lambda(\varphi_c) \ln \left(\frac{M^2}{\varphi_c^2} \right) \right]^2 \right\} .$$

Putting this into V_{eff}, we find

$$V_{\text{eff}}(\varphi_c) = \frac{1}{4!} \left[1 - \frac{25}{64\pi^2} \lambda(\varphi_c) \right] \lambda(\varphi_c) \, \varphi_c^4 .$$

Since $\frac{25}{64\pi^2} \approx 0.04$, we find that for $\lambda(\varphi_c)$ small and positive $V_{\text{eff}}(\varphi_c)$ looks like an ordinary $+\varphi^4$ potential with a modified (but still positive) quartic coupling.

The question now is whether we can trust perturbation theory near $\varphi_c = 0$. Rearranging the solution of the flow equation for λ, we find

$$\lambda(M) = \lambda(\varphi_c) \left[1 + \frac{3}{32\pi^2} \lambda(\varphi_c) \ln \left(\frac{M^2}{\varphi_c^2} \right) \right] .$$

---

[10]This argument follows section V of S. Coleman and E. Weinberg, "Radiative Corrections as the Origin of Spontaneous Symmetry Breaking," Phys. Rev. D, Vol. 7 No. 6, 15 Mar 1973 and Section 18.2 in S. Weinberg, Volume II.

We are interested in taking $\varphi_c \to 0$, so that we may consider M close to but larger than φ_c and thereby take $\ln(M^2/\varphi_c^2) > 0$. For perturbation theory to be valid, we need

$$\left| \frac{3}{32\pi^2} \lambda(\varphi_c) \ln\left(\frac{M^2}{\varphi_c^2}\right) \right| \ll 1$$

to keep the second term in the expansion smaller than the first term.

Assuming that this is true by taking M sufficiently close to φ_c (so that the log is not large) and by taking $\lambda(\varphi_c)$ sufficiently small, then the solution to the flow equation implies that $\lambda(M)$ and $\lambda(\varphi_c)$ have the same sign. For vacuum stability, we take $\lambda(M)$ positive. Therefore, the effective potential is

$$V_{\text{eff}}(\varphi_c) = \frac{1}{4!} \tilde{\lambda}(\varphi_c) \, \varphi_c^4$$

with

$$\tilde{\lambda}(\varphi_c) \equiv \left[1 - \frac{25}{64\pi^2} \lambda(\varphi_c) \right] \lambda(\varphi_c) > 0 \; .$$

The minimum of the effective potential is at $\varphi_c = 0$, and the symmetry $\varphi_c \to -\varphi_c$ of the classical potential is not broken.

For a scalar theory in which a classical symmetry truly is broken by one-loop effects, see problems IV.3.5 and IV.6.9.

Addendum 2: Effective Potential using Dimensional Regularization

It is pedagogically instructive to repeat the calculation of the effective potential using dimensional regularization, in contrast to the cutoff regularization used in the text and in the original paper by Coleman and Weinberg.

Useful formulas:

$$\Gamma(-n + \varepsilon/2) = \frac{(-1)^n}{n!} \left[\frac{2}{\varepsilon} - \gamma + \sum_{k=1}^{n} \frac{1}{k} + O(\varepsilon) \right]$$

$$\int \frac{d^d k_E}{(2\pi)^d} \frac{(k_E^2)^a}{(k_E^2 + D)^b} = \frac{\Gamma(b - a - d/2)\Gamma(a + d/2)}{(4\pi)^{d/2}\Gamma(b)\Gamma(d/2)} D^{-(b-a-d/2)}$$

Metric signature: $\eta = (+, -, -, -)$.

Lagrangian:

$$\mathcal{L} = \frac{1}{2} Z (\partial\varphi)^2 - \frac{1}{2} Z_m m^2 \varphi^2 - \frac{1}{4!} Z_\lambda \lambda \tilde{\mu}^\varepsilon \varphi^4$$

1-loop effective potential:

$$V(\varphi_c) = \frac{1}{2} Z_m m^2 \varphi_c^2 + \frac{1}{4!} Z_\lambda \lambda \tilde{\mu}^\varepsilon \varphi_c^4 + i \int \frac{d^d k}{(2\pi)^d} \sum_{n=1}^{\infty} \frac{1}{2n} \left(\frac{\frac{1}{2}\lambda \tilde{\mu}^\varepsilon \varphi_c^2}{k^2 - m^2 + i\varepsilon} \right)^n$$

Specialize to $d = 4 - \varepsilon \implies \frac{d}{2} = 2 - \frac{\varepsilon}{2}$, $1 - \frac{d}{2} = -1 + \frac{\varepsilon}{2}$, $2 - \frac{d}{2} = +\frac{\varepsilon}{2}$.

Only the $n = 1$ and $n = 2$ integrals diverge, and their divergences will be absorbed in Z_m and Z_λ.

The relevant integrals are the following:

- $n = 1$ (note $k^2 = -k_E^2$ so an overall (-1) gets factored out of the integral):

$$\frac{1}{2}\lambda\tilde{\mu}^\varepsilon\varphi_c^2\int\frac{d^dk}{(2\pi)^d}\frac{1}{k^2-m^2+i\varepsilon} = \frac{1}{2}\lambda\tilde{\mu}^\varepsilon\varphi_c^2(-i)\int\frac{d^dk_E}{(2\pi)^d}\frac{1}{k_E^2+m^2}$$

$$= \frac{1}{2}\lambda\tilde{\mu}^\varepsilon\varphi_c^2(-i)\frac{\Gamma(1-\frac{d}{2})}{(4\pi)^{d/2}}(m^2)^{-(1-\frac{d}{2})}$$

$$= \frac{1}{2}\lambda\tilde{\mu}^\varepsilon\varphi_c^2(-i)\frac{\Gamma(-1+\frac{\varepsilon}{2})}{(4\pi)^2}\left(\frac{4\pi}{m^2}\right)^{\varepsilon/2}m^2$$

$$= \frac{1}{2}\lambda\varphi_c^2(-i)\frac{m^2}{(4\pi)^2}\Gamma(-1+\frac{\varepsilon}{2})\left(\frac{4\pi\tilde{\mu}^2}{m^2}\right)^{\varepsilon/2}$$

$$= \frac{1}{2}\lambda\varphi_c^2(-i)\frac{m^2}{(4\pi)^2}(-1)[\frac{2}{\varepsilon}-\gamma+1+O(\varepsilon)][1+\frac{\varepsilon}{2}\ln\left(\frac{4\pi\tilde{\mu}^2}{m^2}\right)+O(\varepsilon^2)]$$

$$= \frac{1}{2}\lambda\varphi_c^2(+i)\frac{m^2}{(4\pi)^2}[\frac{2}{\varepsilon}+\ln\left(\frac{4\pi e^{-\gamma}\tilde{\mu}^2}{m^2}\right)+1+O(\varepsilon)]$$

So in total we have

$$i\int\frac{d^dk}{(2\pi)^d}\frac{1}{2}\left(\frac{\frac{1}{2}\lambda\tilde{\mu}^\varepsilon\varphi_c^2}{k^2-m^2+i\varepsilon}\right) = -\frac{m^2}{4(4\pi)^2}\lambda\varphi_c^2\left[\frac{2}{\varepsilon}+\ln\left(\frac{4\pi e^{-\gamma}\tilde{\mu}^2}{m^2}\right)+1\right]\ .$$

- $n = 2$ (since the denominator is squared this time, the (-1) doesn't matter):

$$(\frac{1}{2}\lambda\tilde{\mu}^\varepsilon\varphi_c^2)^2\int\frac{d^dk}{(2\pi)^d}\frac{1}{(k^2-m^2+i\varepsilon)^2} = (\frac{1}{2}\lambda\tilde{\mu}^\varepsilon\varphi_c^2)^2(+i)\int\frac{d^dk_E}{(2\pi)^d}\frac{1}{(k_E^2+m^2)^2}$$

$$= (\frac{1}{2}\lambda\tilde{\mu}^\varepsilon\varphi_c^2)^2(+i)\frac{\Gamma(2-\frac{d}{2})}{(4\pi)^{d/2}}(m^2)^{-(2-\frac{d}{2})}$$

$$= (\frac{1}{2}\lambda\tilde{\mu}^\varepsilon\varphi_c^2)^2(+i)\frac{\Gamma(\frac{\varepsilon}{2})}{(4\pi)^2}\left(\frac{4\pi}{m^2}\right)^{\varepsilon/2}$$

$$= \tilde{\mu}^\varepsilon(\frac{1}{2}\lambda\varphi_c^2)^2(+i)\frac{1}{(4\pi)^2}\Gamma(\frac{\varepsilon}{2})\left(\frac{4\pi\tilde{\mu}^2}{m^2}\right)^{\varepsilon/2}$$

$$= \tilde{\mu}^\varepsilon(\frac{1}{2}\lambda\varphi_c^2)^2(+i)\frac{1}{(4\pi)^2}[\frac{2}{\varepsilon}-\gamma+O(\varepsilon)][1+\frac{\varepsilon}{2}\ln\left(\frac{4\pi\tilde{\mu}^2}{m^2}\right)+O(\varepsilon^2)]$$

$$= \tilde{\mu}^\varepsilon(\frac{1}{2}\lambda\varphi_c^2)^2(+i)\frac{1}{(4\pi)^2}[\frac{2}{\varepsilon}+\ln\left(\frac{4\pi e^{-\gamma}\tilde{\mu}^2}{m^2}\right)+O(\varepsilon)]$$

So in total we have

$$i\int\frac{d^dk}{(2\pi)^d}\frac{1}{2\cdot 2}\left(\frac{\frac{1}{2}\lambda\tilde{\mu}^\varepsilon\varphi_c^2}{k^2-m^2+i\varepsilon}\right)^2 = -\frac{\tilde{\mu}^\varepsilon}{16(4\pi)^2}\lambda^2\varphi_c^4\left[\frac{2}{\varepsilon}+\ln\left(\frac{4\pi e^{-\gamma}\tilde{\mu}^2}{m^2}\right)\right]\ .$$

The 1-loop effective potential is (define $\mu^2 \equiv 4\pi \, e^{-\gamma} \tilde{\mu}^2$):

$$V(\varphi_c) =$$

$$\frac{1}{2} Z_m m^2 \varphi_c^2 + \frac{1}{4!} Z_\lambda \lambda \tilde{\mu}^\varepsilon \varphi_c^4 + i \int \frac{d^d k}{(2\pi)^d} \frac{1}{2} \left(\frac{\frac{1}{2}\lambda\tilde{\mu}^\varepsilon \varphi_c^2}{k^2 - m^2 + i\varepsilon} \right) + i \int \frac{d^d k}{(2\pi)^d} \frac{1}{2 \cdot 2} \left(\frac{\frac{1}{2}\lambda\tilde{\mu}^\varepsilon \varphi_c^2}{k^2 - m^2 + i\varepsilon} \right)^2$$

$$+ i \int \frac{d^d k}{(2\pi)^d} \sum_{n=3}^{\infty} \frac{1}{2n} \left(\frac{\frac{1}{2}\lambda\tilde{\mu}^\varepsilon \varphi_c^2}{k^2 - m^2 + i\varepsilon} \right)^n$$

$$= \frac{1}{2} \left\{ Z_m - \frac{\lambda}{64\pi^2} \left[\frac{2}{\varepsilon} + \ln\left(\frac{\mu^2}{m^2}\right) + 1 \right] \right\} m^2 \varphi_c^2 + \frac{1}{4!} \left\{ Z_\lambda - \frac{\lambda}{256\pi^2} \left[\frac{2}{\varepsilon} + \ln\left(\frac{\mu^2}{m^2}\right) \right] \right\} \lambda \tilde{\mu}^\varepsilon \varphi_c^4$$

$$+ i \int \frac{d^d k}{(2\pi)^d} \sum_{n=3}^{\infty} \frac{1}{2n} \left(\frac{\frac{1}{2}\lambda\tilde{\mu}^\varepsilon \varphi_c^2}{k^2 - m^2 + i\varepsilon} \right)^n$$

At this point we are free to take $\tilde{\mu}^\varepsilon \to 1$ everywhere. The regularized 1-loop effective potential is:

$$V_{\text{reg}}(\varphi_c) =$$

$$\frac{1}{2} \left\{ Z_m - \frac{\lambda}{64\pi^2} \left[\frac{2}{\varepsilon} + \ln\left(\frac{\mu^2}{m^2}\right) + 1 \right] \right\} m^2 \varphi_c^2 + \frac{1}{4!} \left\{ Z_\lambda - \frac{\lambda}{256\pi^2} \left[\frac{2}{\varepsilon} + \ln\left(\frac{\mu^2}{m^2}\right) \right] \right\} \lambda \varphi_c^4$$

$$- \int \frac{d^4 k_E}{(2\pi)^4} \sum_{n=3}^{\infty} \frac{(-1)^n}{2n} \left(\frac{\frac{1}{2}\lambda\varphi_c^2}{k_E^2 + m^2} \right)^n$$

We have Wick rotated the integral to Euclidean momentum and set $d = 4$ in all of the convergent integrals. We can now perform the integrals and then resum the series. For each $n \geq 3$, the integral is:

$$\int \frac{d^4 k_E}{(2\pi)^4} \frac{1}{(k_E^2 + m^2)^n} = \frac{1}{(4\pi)^2} \frac{\Gamma(n-2)}{\Gamma(n)} (m^2)^{-(n-2)} = \frac{m^4}{16\pi^2} \frac{1}{(n-1)(n-2)} \frac{1}{(m^2)^n}$$

Thanks to Mathematica, we also have the series:

$$\sum_{n=3}^{\infty} (-1)^n \frac{1}{n(n-1)(n-2)} x^n = \frac{1}{4} \left[x(3x+2) - 2(1+x)^2 \ln(1+x) \right]$$

So defining $x \equiv \lambda\varphi_c^2/(2m^2)$, we have

$$\int \frac{d^4 k_E}{(2\pi)^4} \sum_{n=3}^{\infty} \frac{(-1)^n}{2n} \left(\frac{\frac{1}{2}\lambda\varphi_c^2}{k_E^2 + m^2} \right)^n = \frac{1}{2} \left(\frac{m^4}{16\pi^2} \right) \frac{1}{4} \left[x(3x+2) - 2(1+x)^2 \ln(1+x) \right]$$

$$= \frac{m^4}{128\pi^2} \left[\frac{\lambda\varphi_c^2}{2m^2} \left(\frac{3\lambda\varphi_c^2}{2m^2} + 2 \right) - 2\left(1 + \frac{\lambda\varphi_c^2}{2m^2} \right)^2 \ln\left(1 + \frac{\lambda\varphi_c^2}{2m^2} \right) \right]$$

Therefore, the dimensionally regularized and resummed 1-loop effective potential is (now dropping the subscript on the field):

$$V_{\text{reg}}(\varphi) =$$

$$\frac{1}{2}\left\{Z_m - \frac{\lambda}{64\pi^2}\left[\frac{2}{\varepsilon} + \ln\left(\frac{\mu^2}{m^2}\right) + 1\right]\right\}m^2\varphi^2 + \frac{1}{4!}\left\{Z_\lambda - \frac{\lambda}{256\pi^2}\left[\frac{2}{\varepsilon} + \ln\left(\frac{\mu^2}{m^2}\right)\right]\right\}\lambda\varphi^4$$

$$- \frac{m^4}{128\pi^2}\left[\frac{\lambda\varphi^2}{2m^2}\left(\frac{3\lambda\varphi^2}{2m^2} + 2\right) - 2\left(1 + \frac{\lambda\varphi^2}{2m^2}\right)^2\ln\left(1 + \frac{\lambda\varphi^2}{2m^2}\right)\right]$$

Now we are ready to pick renormalization conditions. We choose the "on-shell" renormalization scheme defined by:

$$\left.\frac{d^2V_{\text{reg}}(\varphi)}{d\varphi^2}\right|_{\varphi=0} \equiv m^2 \ , \quad \left.\frac{d^4V_{\text{reg}}(\varphi)}{d\varphi^4}\right|_{\varphi=\mu} \equiv \lambda(\mu)$$

This is the same as the scheme chosen in Coleman and Weinberg's paper and in the text. Imposing these renormalization conditions fixes:

$$Z_m = 1 + \frac{\lambda}{64\pi^2}\left[\frac{2}{\varepsilon} + 1 + \ln\frac{\mu^2}{m^2}\right]$$

$$Z_\lambda = 1 + \frac{\lambda}{256\pi^2}\left[\frac{2}{\varepsilon} + \ln\frac{\mu^2}{m^2} - 24\ln\left(\frac{\lambda\mu^2}{2m^2} + 1\right) - 48\frac{\mu^2}{m^2}\lambda + O(\lambda^2)\right] \ .$$

Here we are dropping $O(\lambda^3)$ and higher terms, but not Taylor expanding the logs. The reason is that in the $m \to 0$ limit, the $\lambda \ll 1$ in the numerator is offset by $m \to 0$ in the denominator, so that the log is not well approximated by any finite order in its Taylor series.

Putting these back into $V_{\text{reg}}(\varphi)$ gives the renormalized effective potential:

$$V_{\text{ren}}(\varphi) = \frac{1}{24}\left\{\left[\frac{3}{8\pi^2}m^2\ln\left(\frac{\lambda\varphi^2}{2m^2} + 1\right) + 12\varphi^2\right]m^2\right.$$

$$+ \left[\frac{3}{8\pi^2}\left(\ln\left(\frac{\lambda\varphi^2}{2m^2} + 1\right) - \frac{1}{2}\right)m^2\varphi^2 + \varphi^4\right]\lambda$$

$$\left.+ \frac{3}{32\pi^2}\varphi^4\left[\ln\left(\frac{\lambda\varphi^2 + 2m^2}{\lambda\mu^2 + 2m^2}\right) - \frac{3}{2}\right]\lambda^2 + O(\lambda^3)\right\} \ .$$

In the limit $m \to 0$, this becomes:

$$V_{\text{ren}}(\varphi) = \frac{1}{24}\lambda(\mu)\varphi^4 + \frac{\lambda(\mu)^2}{256\pi^2}\varphi^4\left(\ln\frac{\varphi^2}{\mu^2} - \frac{25}{6}\right)$$

which matches equation (20) on p. 242. This is a useful check on the arithmetic: although we have regularized the integrals differently, if we choose the same renormalization scheme then we must get the same answer.

112

IV.4   Magnetic Monopole

2. Show by writing out the components explicitly that $dF = 0$ expresses something that you are familiar with but disguised in a compact notation.

Solution:

$dF = 0 \implies \partial_{[\mu}F_{\nu\rho]} = 0$. (The brackets denote complete antisymmetrization of the indices.) Contract that equation with $\varepsilon^{0\mu\nu\rho}$ and recognize $B^i = \frac{1}{2}\varepsilon^{ijk}F_{jk}$ to get $\vec{\nabla}\cdot\vec{B} = 0$ immediately (Note that choosing one index of $e^{\lambda\mu\nu\rho}$ to be 0 fixes the other indices to be purely spatial.) Now expand $\partial_{[\mu}F_{\nu\rho]} = \frac{1}{3!}\left(\partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} - \partial_\rho F_{\nu\mu} - \partial_\mu F_{\rho\nu} - \partial_\nu F_{\mu\rho}\right) = 0$. Since $F_{\mu\nu} = -F_{\nu\mu}$, the last three terms just combine with the first three terms.

$$\partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} = 0$$

Choose $\mu = i, \rho = j, \nu = 0$ to get

$$\partial_i F_{0j} + \partial_j F_{i0} + \partial_0 F_{ij} = 0$$

The electric field is $E_i = F_{0i}$, so the above is

$$\partial_i E_j - \partial_j E_i + \partial_0 F_{ij} = 0 \implies 2\partial_{[i}E_{j]} + \partial_0 F_{ij} = 0$$

Contract this equation with $\frac{1}{2}\varepsilon^{kij}$ and recognize the definition of the curl of two vectors ($[\vec{a}\times\vec{b}]^i = \varepsilon^{ijk}a_j b_k$) to get

$$\vec{\nabla}\times\vec{E} + \partial_t\vec{B} = 0$$

Introducing the Hodge star operation

$$(^*M)_{\mu_1\ldots\mu_p} = \frac{1}{p!}\varepsilon_{\mu_1\ldots\mu_p\mu_{p+1}\ldots\mu_n}M^{\mu_{p+1}\ldots\mu_D}$$

it is possible to repeat this exercise with $d(^*F) = 0$ as follows. The electromagnetic action on a D-dimensional spacetime $\mathcal{M}$ is:

$$S = \int_{\mathcal{M}}\left(-\frac{1}{4}F\,^*F + A\,^*j\right)$$

A is the 1-form potential, $F = dA$ is its 2-form field strength, and j is a 1-form current that couples to the potential A. If j is a 1-form, then $^*j$ is a $(D-1)$-form, so that the term $A\,^*j$ is a D-form and thus can be integrated over the space M. Varying the action with respect to A as $\delta S \equiv S[A+\delta A] - S[A] = 0$ implies $d(^*F) = \,^*j$. So $d(^*F) = 0$ should be the source-free Maxwell equations $\vec{\nabla}\cdot\vec{E} = 0$ and $\vec{\nabla}\times\vec{B} - \partial_t\vec{E} = 0$.

Now specialize to $D = 4$ and compute:

$$^*F = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}dx^\mu dx^\nu$$

Now take the derivative:
$$d(^*F) = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}\partial_\lambda F^{\rho\sigma}dx^\lambda dx^\mu dx^\nu$$

The * operation is defined such that operating twice gives you back the original thing (hence the name "dual"), so the equation $d(^*F) = {}^*j$ can be written $^*d(^*F) = j$, which is more convenient. The left-hand side is:

$$^*d(^*F) = \frac{1}{2}\varepsilon_{\alpha\mu\nu\lambda}\left(\frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}\partial^\lambda F_{\rho\sigma}\right)dx^\alpha$$

Using
$$\varepsilon_{\mu\nu\alpha\lambda}\varepsilon^{\mu\nu\rho\sigma} = 2(\delta^\rho_\alpha\delta^\sigma_\lambda - \delta^\rho_\lambda\delta^\sigma_\alpha)$$

we get:
$$^*d(^*F) = \frac{1}{2}(\partial^\sigma F_{\alpha\sigma} - \partial^\rho F_{\rho\alpha})dx^\alpha = \partial^\mu F_{\alpha\mu}dx^\alpha$$

So the equation $^*d(^*F) = j$ in components becomes just
$$\partial_\mu F^{\alpha\mu} = j^\alpha$$

Writing the 4-current as $j^\alpha = (\rho, \vec{J})$, setting $\alpha = 0$ and defining again $E^i \equiv F^{0i}$ in the above gives $\partial_i F^{0i} = j^0 \implies \vec{\nabla}\cdot\vec{E} = \rho$. (In verifying this remember $F^{00} = 0$ by antisymmetry.) Now instead of $\alpha = 0$, set $\alpha = i$ to get:

$$\partial_\mu F^{i\mu} = \partial_0 F^{i0} + \partial_j F^{ij} = j^i \implies -\partial_t E^i + \partial_j F^{ij} = J^i$$

Recalling the definition $B^i = \frac{1}{2}\varepsilon^{ijk}F_{jk}$ consider the following:

$$[\vec{\nabla}\times\vec{B}]^k \equiv \varepsilon^{kij}\partial_i B_j = \varepsilon^{kij}\partial_i\left(\frac{1}{2}\varepsilon_{j\ell m}F^{\ell m}\right)$$

$$= \frac{1}{2}(\delta^k_\ell\delta^i_m - \delta^k_m\delta^i_\ell)\partial_i F^{\ell m} = \partial_m F^{km}$$

Therefore the term with the F leftover is exactly curl B, which is what we expect. Therefore, we have
$$\vec{\nabla}\times\vec{B} - \partial_t\vec{E} = \vec{J}$$

$D = 4$ is special because F and $^*F$ are both 2-forms, so there is an electric-magnetic duality that swaps the geometric identity $dF = 0$ with the equation of motion $d(^*F) = 0$.

3. Consider $F = (g/4\pi)d\cos\theta\, d\varphi$. By transforming to Cartesian coordinates show that this describes a magnetic field pointing outward along the radial direction.

Solution:

The coordinate transformation from spherical to cartesian coordinates is

$$x = r\sin\theta\cos\varphi\ ,\quad y = r\sin\theta\sin\varphi\ ,\quad z = r\cos\theta$$

so

$$\cos\theta = \frac{z}{r} = \frac{z}{\sqrt{x^2+y^2+z^2}}\ ,\quad \tan\varphi = \frac{y}{x}$$

$$d\cos\theta = \frac{dz}{r} - \frac{z}{r^2}dr = \frac{1}{r}\left(dz - \frac{z}{r}dr\right)$$

$$dr = d\sqrt{x^2+y^2+z^2} = \frac{1}{\sqrt{x^2+y^2+z^2}}(x\,dx + y\,dy + z\,dz)$$

$$d\cos\theta = \frac{1}{r}\left[dz - \frac{z}{r}\left(\frac{x\,dx + y\,dy + z\,dz}{r}\right)\right] = \frac{1}{r}\left[\left(1 - \frac{z^2}{r^2}\right)dz - \frac{z}{r^2}(x\,dx + y\,dy)\right]$$

$$d\varphi = \frac{1}{x^2+y^2}(-y\,dx + x\,dy)$$

$$d\cos\theta\,d\varphi = \frac{1}{(x^2+y^2)r}\left[\left(1 - \frac{z^2}{r^2}\right)dz(-y\,dx + x\,dy) - \frac{z}{r^2}(x^2dx\,dy - y^2dy\,dx)\right]$$

$$= \frac{1}{(x^2+y^2)r}\left[\left(1 - \frac{z^2}{r^2}\right)dz(-y\,dx + x\,dy) - \frac{z}{r^2}(x^2+y^2)dx\,dy\right]$$

$$1 - \frac{z^2}{r^2} = \frac{1}{r^2}(r^2 - z^2) = \frac{1}{r^2}(x^2+y^2)$$

$$d\cos\theta\,d\varphi = \frac{1}{r^3}[dz(-y\,dx + x\,dy) - z\,dx\,dy]$$

$$dz\,dx = \varepsilon^{312}dy = +dy\ ,\quad dz\,dy = \varepsilon^{321}dx = -dx\ ,\quad dx\,dy = \varepsilon^{123}dz = +dz$$

$$d\cos\theta\,d\varphi = \frac{1}{r^3}(-y\,dy - x\,dx - z\,dz) = -\frac{1}{r^3}(x\,dx + y\,dy + z\,dz)$$

The unit vector in the r direction is precisely $dr = \frac{1}{r}(x\,dx + y\,dy + z\,dz)$, so

$$F = \frac{g}{4\pi}d\cos\theta\,d\varphi = -\frac{g}{4\pi r^2}dr$$

That is a radial magnetic field from a point charge $-g$.

4. Restore the factors of $\hbar$ and c in Dirac's quantization condition.

Solution (due to J. Feinberg):

To do this, we will consider the Lagrangian of a nonrelativistic particle in an electromagnetic

potential. The generalization to a relativistic particle is straightforward and only affects kinetic terms, and that is the first quantized version of field theory. If we are still working in "natural" units, with $\hbar = c = 1$, the Lagrangian is

$$L = \frac{1}{2}m\dot{x}^2 - e\phi(x) + e\dot{\vec{x}} \cdot \vec{A}(\vec{x})$$

with e the charge of the particle. This would give the familiar Lorentz force $e(\vec{E} + \vec{v} \times \vec{B}) \rightarrow e(\vec{E} + (\vec{v}/c) \times \vec{B})$, where we have restored c in Heaviside-Lorentz units (in SI units, we would instead change the dimensions of the magnetic field). Therefore, we need to take the vector potential coupling to

$$e\dot{\vec{x}} \cdot \vec{A}(\vec{x}) \rightarrow \frac{e}{c}\dot{\vec{x}} \cdot \vec{A}(\vec{x}) \ .$$

In the case of the monopole, the needed gauge transformation is

$$\vec{A} \rightarrow \vec{A} - \frac{1}{e}\vec{\nabla}\Lambda, \ \Lambda = \frac{eg}{2\pi}\phi,$$

so our Lagrangian for the particle goes to

$$L \rightarrow L - \frac{1}{c}\dot{\vec{x}} \cdot \vec{\nabla}\Lambda \ .$$

The text states that we must have $e^{i\Lambda}(2\pi) = e^{i\Lambda}(0)$ for the gauge transformation to make sense. Physically, this comes from the fact that the wavefunction of our particle undergoes

$$\psi \rightarrow e^{-i\Lambda}\psi$$

in natural units, and the wavefunction must remain single-valued. We should now check how gauge transformations alter the wavefunction in Heaviside-Lorentz units. Restoring $\hbar$, the amplitude to go from initial wavefunction $\psi_1(\vec{x}_i)$ to a final wavefunction $\psi_2(\vec{x}_f)$ in path integral notation is

$$\langle \psi_2 | \psi_1 \rangle \equiv \int dx_f \psi_2^*(\vec{x}_f) \int dx_i \psi_1(\vec{x}_i) \int_{\vec{x}_i}^{\vec{x}_f} \mathcal{D}x \, e^{\frac{i}{\hbar}\int dt \, L}$$

where L is the one with c restored as discussed, and $\mathcal{D}x$ is the path integral measure. Doing the gauge transformation shifts the phase in the path integral by

$$-\frac{1}{\hbar c}\int dt \, \dot{\vec{x}} \cdot \vec{\nabla}\Lambda = -\frac{1}{\hbar c}[\Lambda(\vec{x}_f) - \Lambda(\vec{x}_i)]$$

because $\dot{\vec{x}} \cdot \vec{\nabla}\Lambda = \partial_t \Lambda(\vec{x}(t))$. The amplitude is therefore left invariant if the wavefunctions transform as

$$\psi \rightarrow e^{-i\Lambda/(\hbar c)}\psi \ .$$

The rest follows as in the text, and we find the condition

$$g = \frac{hcn}{e} \ .$$

where $h = 2\pi\hbar$ as usual. Note also that this is consistent with comparing equations (2) and (3) on p. 307.

5. Write down the reparametrization-invariant current $J^{\mu\nu\lambda}$ of a membrane.

Solution:

The generalization of equation (12) on p. 251 is immediate:

$$J^{\mu\nu\lambda}(x) = \int dX^{\mu} dX^{\nu} dX^{\lambda} \, \delta^{(D)}(x - X) \, .$$

6. Let $g(x)$ be the element of a group G. The 1-form $v = g dg^{\dagger}$ is known as the Cartan-Maurer form. Then $\mathrm{tr} v^N$ is trivially closed on an N-dimensional manifold since it is already an N-form. Consider $Q = \int_{S^N} \mathrm{tr} v^N$ with S^N the N-dimensional sphere. Discuss the topological meaning of Q. These considerations will become important later when we discuss topology in field theory in chapter V.7. [Hint: Study the case $N = 3$ and $G = SU(2)$.]

Solution:

Let's consider the case for which G is some group whose manifold has no pathologies, for instance S^N, so that infinitesimal deviations from any point in G are sufficient to determine the global structure of object Q. So consider letting $g \to g + \delta g$. Then v changes as

$$v \to (g + \delta g)d(g^{-1} + \delta g^{-1}) = v + \delta g dg^{-1} + g d(\delta g^{-1}) + O(\delta g^2)$$

Since $gg^{-1} = I$ (the identity element on G), we have:

$$(g + \delta g)(g^{-1} + \delta g^{-1}) = gg^{-1} + \delta g g^{-1} + g\delta g^{-1} + O(\delta g^2) = I \implies \delta g^{-1} = -g^{-1}\delta g \, g^{-1}$$

So v transforms as

$$
\begin{aligned}
v \to v &+ \delta g \, dg^{-1} - g d(g^{-1}\delta g \, g^{-1}) \\
&= \delta g \, dg^{-1} - \left[d(\delta g \, g^{-1}) - dg \, g^{-1}\delta g \, g^{-1}\right] \\
&= \delta g \, dg^{-1} - \left[d(\delta g)g^{-1} + \delta g \, dg^{-1} - dg \, g^{-1}\delta g \, g^{-1}\right] \\
&= -d(\delta g)g^{-1} + dg \, g^{-1}\delta g \, g^{-1} \\
&= -d(gg^{-1}\delta g)g^{-1} + dg \, g^{-1}\delta g \, g^{-1} \\
&= -g \, d(g^{-1}\delta g)g^{-1}
\end{aligned}
$$

117

So under $g \to g + \delta g$, we have $\delta v \equiv v[g + \delta g] - v[g] = -g\, d(g^{-1}\delta g)g^{-1}$, and therefore:

$$\delta Q = \int \mathrm{tr}\; \delta v\, v...v + \int \mathrm{tr}\; v\delta v...v + ... + \int \mathrm{tr}\; vv...\delta v = N \int \mathrm{tr}\; vv...v\,\delta v$$

$$= -N \int \mathrm{tr}\; vv...v\, gd(g^{-1}\delta g)g^{-1}$$

$$= -N \int \mathrm{tr}\; (gdg^{-1})(gdg^{-1})...(gdg^{-1})g\, d(g^{-1}\delta g)g^{-1}$$

$$= -N \int \mathrm{tr}\; dg^{-1}\, gdg^{-1}...gdg^{-1}\, gd(g^{-1}\delta g) \quad \text{(by cyclicity of trace)}$$

$$= +N \int \mathrm{tr}\; dg^{-1}g\, dg^{-1}g\,...gdg^{-1}\,(gg^{-1})dg\, d(g^{-1}\delta g) \quad \text{(by } dg^{-1}g = d(g^{-1}g) - g^{-1}dg = 0 - g^{-1}dg)$$

$$= +N \int \mathrm{tr}\; dg^{-1}g\, dg^{-1}g\,...gdg^{-1}\, dg\, d(g^{-1}\delta g) \quad \text{(since } g^{-1}g = I)$$

$$= (-1)^{N-1}N \int \mathrm{tr}\; dg\, dg\,...\, dg\, d(g^{-1}\delta g) \quad \text{(by repeating the above another } (N-1) \text{ times)}$$

$$= (-1)^{N-1}N \int d\left[\mathrm{tr}\; dg\, dg\,...\, dg\, g^{-1}\delta g\right] \quad \text{(since } d^2 = 0)$$

$$= 0$$

So the quantity Q calculated at a point g on the group space G and the quantity Q calculated at a point $g + \delta g$ on G are the same. Therefore, barring any unforeseen pathologies on the space G, we can compound the infinitesimal transformations and conclude that we can calculate Q using any point on the group manifold and get the same answer. Thus Q is a topological quantity, which depends only on the particular group G we pick. Now that we know Q is topological, let's try to figure out what it means.

First consider the case $G = U(1)$ and $N = 1$. Then $g(x) = e^{in\theta(x)}$, so $v = gdg^{-1} = ne^{i\theta}d(e^{-i\theta}) = -nid\theta$, and the Cartan-Maurer form is

$$Q = \int_{S^1} v = -in \int_{S^1} d\theta = -i2\pi n \;, \quad n \in \mathbb{Z}$$

In conclusion, for this case the quantity $Q/(-2\pi i)$ counts the number of times the spatial circle wraps around the group circle:

$$\frac{Q}{-2\pi i} \; \in \; \Pi_1(S^1)$$

For any N, the object Q properly normalized counts the number of times the spatial N-sphere wraps around the group manifold G:

$$Q \; \in \; \Pi_N(G)$$

where Q is suitably normalized. In particular, for $G = S^N$, $\Pi_N(S^N) = \mathbb{Z}$. This mathematical fact, that Q is proportional to an integer determined purely by topology, tells us that the chiral anomaly does not get renormalized (see Chapter IV.7).

IV.5 Nonabelian Gauge Theory

3. In 4 dimensions $\varepsilon^{\mu\nu\lambda\rho}\mathrm{tr}F_{\mu\nu}F_{\lambda\rho}$ can be written as $\mathrm{tr}F^2$. Show that $d\mathrm{tr}F^2 = 0$ in any dimensions.

Solution:

Using differential forms notation, we have

$$d\,\mathrm{tr}F^2 = \mathrm{tr}(dFF + (-1)^2 FdF) = 2\,\mathrm{tr}\,dFF \quad \text{(cyclic property of trace)}$$
$$= -2\,\mathrm{tr}[A, F]F = 0$$

where in the last step we have used the Bianchi identity $DF \equiv dF + [A, F] = 0$.

This result is equation (A.15) in B. Zumino, Wu Yong-Shi, A. Zee, "Chiral Anomalies, Higher Dimensions, and Differential Geometry," Nucl. Phys. B239:477 (1984)

5. For a challenge show that $\mathrm{tr}F^n$, which appears in higher dimensional theories such as string theory, are all total divergences. In other words, there exists a $(2n-1)$-form $\omega_{2n-1}(A)$ such that $\mathrm{tr}F^n = d\omega_{2n-1}(A)$. [Hint: A compact representation of the form $\omega_{2n-1}(A) = \int_0^1 dt\, f_{2n-1}(t, A)$ exists.] Work out $\omega_5(A)$ explicitly and try to generalize knowing ω_3 and ω_5. Determine the $(2n-1)$-form $f_{2n-1}(t, A)$. For help, see B. Zumino et al., *Nucl. Phys.* B239:477, 1984.

Solution:

The Bianchi identity, $DF \equiv dF + [A, F] = 0$ will be useful.

$$d\,(\mathrm{tr}\,F^n) = n\,\mathrm{tr}\,dF\,F^{n-1}$$
$$= -n\,\mathrm{tr}\,[A, F]F^{n-1} \qquad \text{(Bianchi identity)}$$
$$= -n\,\mathrm{tr}\,\left(AF^n - FAF^{n-1}\right)$$
$$= -n\,\mathrm{tr}\,\left(AF^n - AF^n\right) \qquad \text{(cyclicity of trace)}$$
$$= 0$$

Therefore $\mathrm{tr}\,F^n$ is locally a total divergence:

$$\mathrm{tr}\,F^n = d(\text{something}) \equiv d\omega_{2n-1}(A)$$

Now let us find this $(2n-1)$-form $\omega_{2n-1}(A)$. We follow the reference B. Zumino, Wu Yong-Shi, A. Zee, "Chiral Anomalies, Higher Dimensions, and Differential Geometry," Nucl. Phys. B239:477 (1984).

Given a Yang-Mills potential A, define a 1-parameter class of potentials $A_s \equiv sA$ and their associated field strength tensors $F_s = dA_s + A_s^2 = s\, dA + s^2 A^2$. Now differentiate:

$$\frac{d}{ds}(\operatorname{tr} F_s^n) = n\operatorname{tr}\left(\frac{dF_s}{ds}F_s^{n-1}\right)$$

$$= n\operatorname{tr}\left(\frac{d}{ds}(s\, dA + s^2 A^2)F_s^{n-1}\right)$$

$$= n\operatorname{tr}\left((dA + 2sA^2)F_s^{n-1}\right)$$

$$= n\operatorname{tr}\left((dA + 2A_sA)F_s^{n-1}\right)$$

The commutator of a Yang-Mills 1-form A_s with a p-form X is:

$$[A_s, X] = A_sX - (-1)^p X A_s$$

So letting $X = A$ and thus $p = 1$ gives

$$[A_s, A] = A_sA - (-1)AA_s = A_sA + AA_s$$

But $A_s = sA$, so I can take the s from one A and put it in the other A to give

$$[A_s, A] = 2A_sA$$

So we have

$$\frac{d}{ds}(\operatorname{tr} F_s^n) = n\operatorname{tr}\left((dA + [A_s, A])F_s^{n-1}\right)$$

$$= n\operatorname{tr}\left(D_sA\, F_s^{n-1}\right)$$

$D_s = d + [A_s, \,\cdot\,]$ is the covariant derivative associated with the potential A_s. Using the Bianchi identity $D_sF_s = 0$, the covariant derivative can act on the product AF_s^{n-1}:

$$\frac{d}{ds}(\operatorname{tr} F_s^n) = n\operatorname{tr} D_s\left(A\, F_s^{n-1}\right)$$

$$= n\, d\operatorname{tr} AF_s^{n-1} + n\operatorname{tr}\left[A_s,\, AF_s^{n-1}\right]$$

F_s is a 2-form, so F_s^{n-1} is a $2(n-1)$-form. A is a 1-form, so AF_s^{n-1} is a $(2n-1)$-form, which is odd for all n. So the commutator gives a *plus* sign:

$$\operatorname{tr}\left[A_s,\, AF_s^{n-1}\right] = \operatorname{tr}\left(A_sAF_s^{n-1} + AF_s^{n-1}A_s\right)$$

It is tempting to use the cyclicity of the trace to have those two terms add, but we have to be careful. Let Ω be a $(2n-1)$-form. Consider the following trace:

$$\operatorname{tr}(\Omega A) = \operatorname{tr}(\Omega_{\mu_1\ldots\mu_{2n-1}}A_\mu)\, dx^{\mu_1}\ldots dx^{\mu_{2n-1}}dx^\mu$$

Cyclicity of the trace indeed lets us move the matrix A_μ to the left of the matrix $\Omega_{\mu_1\ldots\mu_{2n-1}}$. But to repackage the expression into forms notation, we have to move the dx^μ to the left of all of the other dx^{μ_i}, with each exchange picking up a minus sign. Since there are $(2n-1)$

exchanges, we pick up an overall minus sign. So $\operatorname{tr}(\Omega A) = -\operatorname{tr}(A\Omega)$, for Ω being an arbitrary $(2n-1)$-form.

Since AF_s^{n-1} is also a $(2n-1)$-form, we have $\operatorname{tr}[A_s, AF_s^{n-1}] = 0$. Therefore:

$$\frac{d}{ds}(\operatorname{tr} F_s^n) = n\, d\operatorname{tr} AF_s^{n-1}$$

Now, remember what we want is $\operatorname{tr} F^n = \lim_{s\to 1} \operatorname{tr} F_s^n$, and note that $\lim_{s\to 0} \operatorname{tr} F_s^n = 0$. So if we integrate the above expression from $s = 0$ to 1, we get the desired result:

$$\operatorname{tr} F^n = d\omega_{2n-1}(A)\,, \quad \omega_{2n-1}(A) = n\int_0^1 ds\, \operatorname{tr} A(s\,dA + s^2 A^2)^{n-1}$$

Now to figure out what all this means, recall question IV.4.6 whose result was that, for some element g of the group G, the object $Q = \int_{S^N} \operatorname{tr}(g^{-1}dg)$ suitably normalized is an element of the homotopy class $\Pi_N(G)$.

If $A(x \,\epsilon\, S^N) = g^{-1}dg$ (that is, if A approaches the gauge transformation of 0 on the spatial S^N) then:

$$dA = d(g^{-1}dg) = dg^{-1}dg$$

and

$$AA = g^{-1}dg\, g^{-1}dg = -dg^{-1}gg^{-1}dg = -dg^{-1}dg$$

So $dA = -A^2 = -(g^{-1}dg)^2$. Therefore, we have $dA + sA^2 = -(1-s)A^2$ and

$$\omega_{2n-1}(A) = (-1)^{n-1}n\int_0^1 ds\, s^{n-1}(1-s)^{n-1}\operatorname{tr} A^{2n-1}$$

Meanwhile, the integral over the parameter s evaluates to

$$\int_0^1 ds\, s^{n-1}(1-s)^{n-1} = \frac{\pi^{1/2}}{2^{2n-1}}\frac{\Gamma(n)}{\Gamma(n+1/2)}$$

So remembering that $n\Gamma(n) = \Gamma(n+1)$, we have the result:

$$\omega_{2n-1}(A) = (-1)^{n-1}\frac{\pi^{1/2}}{2^{2n-1}}\frac{\Gamma(n+1)}{\Gamma(n+1/2)}\operatorname{tr}(g^{-1}dg)^{2n-1}$$

Thus for $g \,\epsilon\, G$, and for $2n$-dimensional Euclidean space ($\mathbb{E}^{2n}$) whose boundary can be taken as spherical, we have the result:

$$\int_{\mathbb{E}^{2n}} \operatorname{tr} F^n = \int_{\mathbb{E}^{2n}} d\omega_{2n-1} = \int_{S^{2n-1}} \omega_{2n-1} \propto \int_{S^{2n-1}} \operatorname{tr}(g^{-1}dg)^{2n-1} \,\epsilon\, \Pi_{2n-1}(G)$$

The quantity on the left is the nonabelian generalization of the chiral anomaly, and the quantity on the right is an integer determined purely by topology.

IV.6 Anderson-Higgs Mechanism

1. Consider an $SU(5)$ gauge theory with a Higgs field φ transforming as the 5-dimensional representation: φ^i, $i = 1, ..., 5$. Show that a vacuum expectation value of φ breaks $SU(5)$ to $SU(4)$. Now add another Higgs field φ', also transforming as the 5-dimensional representation. Show that the symmetry can either remain at $SU(4)$ or be broken to $SU(3)$.

Solution:

Consider an $SU(N)$ gauge theory with a Higgs field φ in the N-dimensional ("defining") representation. $SU(N)$ is the collection of transformations that leave the norm

$$(\varphi, \varphi)_N \equiv \sum_{i=1}^{N} \varphi^{\dagger i} \varphi_i$$

invariant. Suppose that the $SU(N)$ symmetry is broken by φ obtaining a vacuum expectation value (VEV). We can always choose a coordinate system (in field space) such that this VEV points in the Nth direction: $\langle \varphi \rangle = (0, ..., 0, v)$. Writing the field as a fluctuation about this classical value

$$\varphi_i = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{N-1} \\ v + \varphi_N \end{pmatrix}$$

changes the norm (φ, φ) into:

$$(\varphi, \varphi)_N = \sum_{i=1}^{N-1} \varphi^{\dagger i} \varphi_i + (v + \varphi_N)^{\dagger}(v + \varphi_N)$$

If $v = 0$, then this just gets repackaged into the sum $\sum_{i=1}^{N} \varphi^{\dagger i} \varphi_i$ as before. If $v \neq 0$, then there is a term linear in φ_N, which is clearly not invariant under a full $SU(N)$ transformation. However, the piece $(\varphi, \varphi)_{N-1} \equiv \sum_{i=1}^{N-1} \varphi^{\dagger i} \varphi_i$ is invariant under $SU(N-1)$ transformations, since it is precisely the $SU(N-1)$-invariant norm. Since the $SU(N)$-invariant Lagrangian (before spontaneous symmetry breaking) depends only on the $SU(N)$ norm $(\varphi, \varphi)_N$, after spontaneous symmetry breaking the Lagrangian's dependence on $\{\varphi_i\}_{i=1}^{N-1}$ will be completely in terms of the $SU(N-1)$ norm $(\varphi, \varphi)_{N-1}$. Therefore, the Higgs VEV breaks $SU(N)$ to $SU(N-1)$.

Now consider an $SU(N)$ gauge theory with two Higgs fields in the N-dimensional representation. $SU(N)$ is the collection of transformations that leave the norm

$$(\varphi, \varphi')_N \equiv \sum_{i=1}^{N} \varphi^{\dagger i} \varphi'_i$$

invariant. Suppose again that the $SU(N)$ symmetry is broken by φ obtaining a vacuum expectation value (VEV). As before, we can choose a coordinate system in field space such that this VEV points in the Nth direction: $\langle\varphi\rangle = (0,...,0,v)$. The question is, in which direction does the VEV of the other Higgs field, φ', point? This direction is in principle given to us by the potential.

If this VEV also happens to be in the Nth direction, then $\langle\varphi'\rangle = (0,...,0,v')$. If this VEV happens to be perpendicular to the Nth direction, then we can still choose coordinates such that $\langle\varphi'\rangle$ is aligned along one of the other directions, say the $(N-1)^{\text{th}}$ direction: $\langle\varphi'\rangle = (0,...,0,v',0)$, where now there are $(N-2)$ zeros to the left of v' rather than $(N-1)$ zeros as in the previous case. It may be the case that this VEV is somewhere in between, neither parallel nor perpendicular to the first VEV. We therefore parameterize the most general case as:

$$
\varphi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{N-1} \\ v + \varphi_N \end{pmatrix} \quad , \quad \varphi' = \begin{pmatrix} \varphi'_1 \\ \varphi'_2 \\ \vdots \\ \varphi'_{N-2} \\ v'\cos\beta + \varphi'_{N-1} \\ v'\sin\beta + \varphi'_N \end{pmatrix}
$$

When $\beta = 0$, the second VEV is perpendicular to the first, and when $\beta = \pi/2$, the second VEV is parallel to the first. Putting this parameterization into the $SU(N)$-invariant norm gives

$$
\begin{aligned}
(\varphi, \varphi')_N &= \sum_{i=1}^{N-2} \varphi^{\dagger\,i}\varphi'_i + \varphi^{\dagger}_{N-1}(v'\cos\beta + \varphi'_{N-1}) + (v+\varphi_N)^{\dagger}(v'\sin\beta + \varphi'_N) \\
&= (\varphi,\varphi')_{N-2} + \varphi^{\dagger}_{N-1}\varphi'_{N-1} + \varphi^{\dagger}_N\varphi'_N + v'\cos\beta\,\varphi^{\dagger}_{N-1} + v^{\dagger}\varphi'_N + v'\sin\beta\varphi^{\dagger}_N + v^{\dagger}v'\sin\beta
\end{aligned}
$$

For generic $\beta \neq \pi/2$, there are terms linear in φ_{N-1} and in φ_N, which are not invariant under $SU(N)$ or $SU(N-1)$ transformations; this expression is only invariant under $SU(N-2)$, as expressed by the $SU(N-2)$-invariant norm $(\varphi,\varphi')_{N-2}$. For the special case $\beta = \pi/2$, the term linear in φ_{N-1} drops out, so we can repackage the φ_{N-1} into an $SU(N-1)$-invariant norm.

So for the particular case $N = 5$, having two Higgs fields transforming as a 5 generically breaks $SU(5)$ to $SU(3)$. In the special case for which the two VEVs happen to be aligned, $SU(5)$ instead breaks to $SU(4)$.

2. In general, there may be several Higgs fields belonging to various representations labeled by α. Show that the mass squared matrix for the gauge bosons generalizes immediately to $(\mu^2)^{ab} = \sum_\alpha g^2 (T_\alpha^a v_\alpha \cdot T_\alpha^b v_\alpha)$, where v_α is the vacuum expectation value of φ_α and T_α^a is the ath generator represented on φ_α. Combine the situations described in exercises IV.6.1 and IV.6.2 and work out the mass spectrum of the gauge bosons.

Solution:

Let φ transform under the R-representation of some gauge group G. Then $D_\mu \varphi = \partial_\mu \varphi - ig A_\mu \varphi$, where $A_\mu = \sum_{a=1}^{\dim G} A_\mu^a T_R^a$ is the matrix-valued gauge field, and T_R^a is the ath Lie algebra generator of the R-representation of the group G. The Lagrangian $\mathcal{L} = (D_\mu \varphi)^\dagger D^\mu \varphi = g^2 \varphi^\dagger A A \varphi + \ldots = \sum_{a,b} g^2 \varphi^\dagger T_R^a T_R^b \varphi \, A_\mu^a A^{b\mu} + \ldots = \sum_{a,b} \frac{1}{2} g^2 \varphi^\dagger \{T_R^a, T_R^b\} \varphi \, A_\mu^a A^{b\mu} + \ldots$ implies, upon spontaneous symmetry breaking $\varphi = V + \ldots$ the gauge boson mass-squared matrix

$$(\mu^2)^{ab} = g^2 V^\dagger \{T_R^a, T_R^b\} V$$

We can also write the complex vector V as $V = \frac{1}{\sqrt{2}} v \, e^{i\theta}$ to get:

$$(\mu^2)^{ab} = \frac{1}{2} g^2 v^T \{T_R^a, T_R^b\} v$$

Now suppose there are $\alpha = 1, \ldots, N$ Higgs fields, each of which transforms under a representation R_α of the group G. Each has its own covariant derivative $D_\mu \varphi_\alpha = \partial_\mu \varphi_\alpha - ig A_\mu^a T_{R_\alpha}^a \varphi_\alpha$, so the kinetic term $\mathcal{L} = (D\varphi)^\dagger D\varphi$ yields

$$\mathcal{L} = \sum_{\alpha=1}^{N} (D\varphi_\alpha)^\dagger D\varphi_\alpha = \sum_{\alpha=1}^{N} \sum_{a,b=1}^{\dim G} g^2 V_\alpha^\dagger T_{R_\alpha}^a T_{R_\alpha}^b V_\alpha \, A_\mu^a A^{b\mu} + \ldots$$

Immediately we have the mass-squared matrix

$$(\mu^2)^{ab} = g^2 \sum_{\alpha=1}^{N} V_\alpha^\dagger \{T_{R_\alpha}^a, T_{R_\alpha}^b\} V_\alpha$$

Now let's specialize to the situation from problem IV.6.1. We have $N = 2$ Higgs fields, both of which transform under the 5-representation of $G = SU(5)$. [The group $SU(n)$ has $n^2 - 1$ generators, so $\dim G = 5^2 - 1 = 24$.] Taking the vacuum expectation value of the first field φ to be $\langle \varphi \rangle = V(0,0,0,0,1)^T$, we can without loss of generality take the vacuum expectation value of the second field φ' to lie in the (φ_4, φ_5)-plane:

$$\langle \varphi' \rangle = V' \begin{pmatrix} 0 \\ 0 \\ 0 \\ \cos\beta \\ \sin\beta \end{pmatrix}$$

This gives

$$(\mu^2)^{ab} = g^2 \left[c^2 V'^2 (t^{ab})_{44} + 2cs V'^2 (t^{ab})_{45} + (V^2 + s^2 V'^2)(t^{ab})_{55} \right]$$

124

where we have defined $c \equiv \cos \beta$, $s \equiv \sin \beta$ and a set of symmetric matrices t^{ab} with components:

$$(t^{ab})_{ij} \equiv \{T_5^a, T_5^b\}_{ij}$$

where T_5^a is the ath generator of the 5-dimensional representation of $SU(5)$.

As explained in IV.6.1, the $SU(5)$ symmetry can be broken down to either $SU(4)$ or $SU(3)$, depending on whether $\langle \varphi \rangle$ and $\langle \varphi' \rangle$ point along the same axis. The above mass matrix is another manifestation of that statement; we see that if $\cos \beta = 0$, then

$$(\mu^2)^{ab} = g^2(V^2 + V'^2)(t^{ab})_{55}$$

so that whichever gauge bosons were massless with just the first Higgs field remain massless with the addition of the second field as long as the two vacuum expectation values point in the same direction. Alternatively, if $\sin \beta = 0$ then

$$(\mu^2)^{ab} = g^2 \left[V'^2(t^{ab})_{44} + V^2(t^{ab})_{55} \right]$$

which shows that if the vacuum expectation values are perpendicular, each Higgs breaks a separate direction in $SU(5)$ and thereby gives mass to the corresponding gauge bosons.

4. In chapter IV.5 you worked out an $SU(2)$ gauge theory with a scalar field φ in the $I = 2$ representation. Write down the most general quartic potential $V(\varphi)$ and study the possible symmetry breaking patterns.

Solution:

The "$I = 1$" (or "spin-1") representation of $SU(2)$ is the symmetric 2×2 matrix, which has 3 components and thus can be written as a 3-component vector under $SO(3)$. The "$I = 2$" (or "spin-2") representation of $SU(2)$ can be written as a symmetric traceless 3×3 tensor under $SO(3)$. Let $i = 1, 2, 3$ denote the 3-vector under $SO(3)$. Then the object $\varphi_{ij} \equiv \varphi_{(ij)} - \frac{1}{3}\text{tr}(\varphi)\delta_{ij}$ transforms under the "spin-2" representation of $SU(2)$. The most general renormalizable potential for φ in $d = 3 + 1$ spacetime dimensions is

$$V(\varphi) = \frac{1}{2}m^2\text{tr}(\varphi^2) + \lambda[\text{tr}(\varphi^2)]^2 + \lambda'\text{tr}(\varphi^4) \ .$$

But you have already encountered this potential way back in problem I.10.3, whose solution (in the book) points out that $\text{tr}(\varphi^4)$ and $[\text{tr}(\varphi^2)]^2$ are actually proportional to each other. We can therefore set $\lambda' = 0$ with no loss of generality, and thus observe that the potential actually has an accidental $SO(5)$ symmetry. We can repackage the five independent numbers φ_{11}, φ_{12}, φ_{13}, φ_{22}, φ_{23} into a 5-dimensional column vector $\vec{\phi} = (\phi_1, ..., \phi_5)^T$, where we will not actually display the explicit relations between the $\{\phi_A\}_{A=1}^5$ and the φ_{ij}. The potential can be rewritten as

$$V(\vec{\phi}) = \frac{1}{2}m^2\vec{\phi} \cdot \vec{\phi} + \lambda(\vec{\phi} \cdot \vec{\phi})^2$$

which is manifestly invariant under the $SO(5)$ transformation $\vec{\phi} \to M\vec{\phi}$, with M any 5-by-5 orthogonal matrix.

Suppose that $m^2 < 0$ and $\lambda > 0$ so that the potential exhibits spontaneous symmetry breaking. By $SO(5)$ invariance, we can choose field coordinates for which the vacuum expectation value of $\vec{\phi}$ points purely in the fifth direction: $\langle\vec{\phi}\rangle = v(0,0,0,0,1)^T$. To study small oscillations about this vacuum, write

$$\vec{\phi} = \begin{pmatrix} \chi_1 \\ \chi_2 \\ \chi_3 \\ \chi_4 \\ v + H \end{pmatrix}$$

and study the $SO(5)$ invariant norm

$$\vec{\phi} \cdot \vec{\phi} = \vec{\chi} \cdot \vec{\chi} + (v + H)^2$$

where the vector arrow over χ runs over only 4 indices, $\vec{\chi} = (\chi_1, ..., \chi_4)^T$. The analysis is just as in IV.6.1: if $v \neq 0$, then the norm $\vec{\phi} \cdot \vec{\phi}$ is invariant only under $SO(4)$ transformations on the $\vec{\chi}$ fields. With only one scalar field in the theory, any negative mass squared instability will break $SO(5)$ to $SO(4)$.

The situation is more complicated with additional scalar fields. If we stick with $SO(5)$, meaning introduce $\vec{\phi}$ and $\vec{\phi}'$ both transforming as 5-vectors under $SO(5)$, then the analysis proceeds analogously to that in problem IV.6.1: the $SO(5)$ symmetry can break either to $SO(4)$ or to $SO(3)$. If instead we insist on starting with traceless symmetric tensors φ_{ij} and φ'_{ij} of $SO(3)$, then we have to worry about cross terms of the form

$$V(\varphi, \varphi') = \lambda_1 [\text{tr}(\varphi^2)][\text{tr}(\varphi'^2)] + \lambda_2 \text{tr}(\varphi^2 \varphi'^2) + \lambda_3 \text{tr}(\varphi\varphi'\varphi\varphi') + \lambda_4 \text{tr}(\varphi^3\varphi) + \lambda_5 \text{tr}(\varphi\varphi'^3)$$

in which case the analysis is more complicated.

5. Complete the derivation of the Feynman rules for the theory in (3) and compute the amplitude for the physical process $\chi + \chi \to B + B$.

Solution:

The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \frac{1}{2}M^2 B_\mu B^\mu + \frac{1}{2}(\partial\chi)^2 - \frac{1}{2}m^2\chi^2 + \mathcal{L}_{\text{int}}$$

where $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$, $M = ev$, $m = \sqrt{2\lambda}\,v$ and

$$\mathcal{L}_{\text{int}} = eM\chi B_\mu B^\mu + \frac{1}{2}e^2\chi^2 B_\mu B^\mu - \lambda v\chi^3 - \frac{1}{4}\lambda\chi^4$$

126
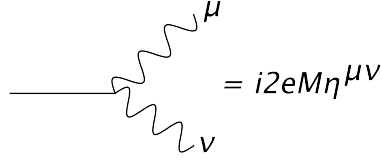
where we have dropped an additive constant. The propagator for χ (solid line) is

$$i\Delta(q) = \frac{i}{q^2 - m^2 + i\varepsilon}$$

and the propagator for the massive photon B_μ (wavy line) is
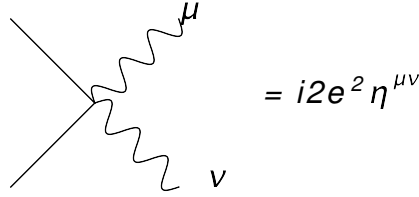
$$i\Delta_{\mu\nu}(q) = \frac{i}{q^2 - M^2 + i\varepsilon}\left(-\eta_{\mu\nu} + \frac{q_\mu q_\nu}{M^2}\right) \; .$$

The cubic χBB vertex comes from $\mathcal{L} = eM\chi B^2 = \frac{1}{2}(2eM\eta^{\mu\nu})\chi B_\mu B_\nu$ and is represented diagrammatically as:
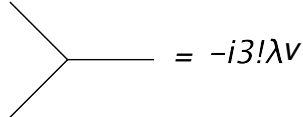
$$= i2eM\eta^{\mu\nu}$$
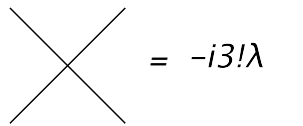
Similarly, the quartic $\chi\chi BB$ vertex comes from $\mathcal{L} = \frac{1}{2}e^2\chi^2 B^2 = \frac{1}{4}(2e^2\eta^{\mu\nu})\chi^2 B_\mu B_\nu$ and is represented diagrammatically as:

$$= i2e^2\eta^{\mu\nu}$$
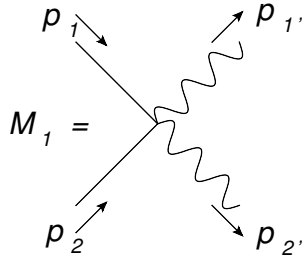
The cubic and quartic self-couplings for χ come from $\mathcal{L} = -\lambda v\chi^3 - \frac{1}{4}\lambda\chi^4 = \frac{1}{3!}(-3!\lambda v)\chi^3 + \frac{1}{4!}(-3!\lambda)\chi^4$, which are represented diagrammatically as:

$$= -i3!\lambda v$$

$$= -i3!\lambda$$

The tree-level amplitude for $\chi\chi \to BB$ involves four diagrams: $\mathcal{M} = \sum_{i=1}^{4}\mathcal{M}_i$, where the $\mathcal{M}_i$ are defined as follows. The first, $\mathcal{M}_1$, is the contribution from the quartic $\chi^2 B^2$ vertex:

This gives
$$\mathcal{M}_1 = \varepsilon_{1'\mu}\varepsilon_{2'\nu}(i2e^2)\eta^{\mu\nu} .$$

Next we have a contribution from s-channel exchange of a χ:
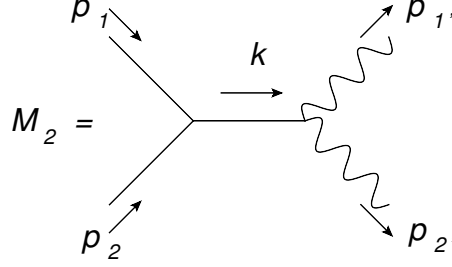
This gives
$$\mathcal{M}_2 = \varepsilon_{1'\mu}\varepsilon_{2'\nu}\left[(-i6\lambda v)i\Delta(k)(+i2e^2 v\eta^{\mu\nu})\right]$$
$$= i\varepsilon_{1'\mu}\varepsilon_{2'\nu}(12\lambda M^2)\frac{1}{s-M^2}\eta^{\mu\nu}$$

where we have defined $k = p_1 + p_2 = p_{1'} + p_{2'}$ and $s \equiv k^2$, and we have used $M = ev$.

The third contribution to the amplitude is from a t-channel exchange of a B_μ:
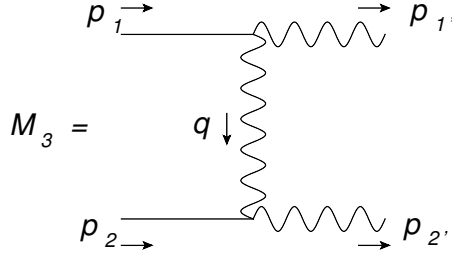
This gives
$$\mathcal{M}_3 = \varepsilon_{1'\mu}\varepsilon_{2'\nu}\left[(i2eM\eta^{\mu\rho})i\Delta_{\rho\sigma}(q)(i2eM\eta^{\sigma\nu})\right]$$
$$= i\varepsilon_{1'\mu}\varepsilon_{2'\nu}(4e^2M^2)\frac{1}{t-M^2}\left(\eta^{\mu\nu} - \frac{q^\mu q^\nu}{M^2}\right)$$

where $q = p_1 - p_{1'} = p_{2'} - p_2$ and $t \equiv q^2 = (p_1 - p_{1'})^2 = (p_{2'} - p_2)^2$. We can use gauge invariance in the form of $\varepsilon_{1'}\cdot p_{1'} = 0$ and $\varepsilon_{2'}\cdot p_{2'} = 0$ to write $\varepsilon_{1'}\cdot q = \varepsilon_{1'}\cdot p_1$ and $\varepsilon_{2'}\cdot q = -\varepsilon_{2'}\cdot p_2$, so that this part of the amplitude becomes
$$\mathcal{M}_3 = i\varepsilon_{1'\mu}\varepsilon_{2'\nu}(4e^2M^2)\frac{1}{t-M^2}\left(\eta^{\mu\nu} + \frac{p_1^\mu p_2^\nu}{M^2}\right) .$$

The fourth is the crossed version of $\mathcal{M}_3$, which results in
$$\mathcal{M}_4 = i\varepsilon_{1'\mu}\varepsilon_{2'\nu}(4e^2M^2)\frac{1}{u-M^2}\left(\eta^{\mu\nu} + \frac{p_2^\mu p_1^\nu}{M^2}\right)$$

128

where $u \equiv (p_{2'} - p_1)^2 = (p_2 - p_{1'})^2$.

The full amplitude $\mathcal{M} = \sum_{i=1}^4 \mathcal{M}_i$ is therefore (recall $v = M/e$):

$$\mathcal{M} = i\varepsilon_{1'\mu}\varepsilon_{2'\nu} \left\{ 2e^2 \left(1 + \frac{6\lambda v^2}{s - M^2} \right) \eta^{\mu\nu} + (4e^2 M^2) \left[\frac{1}{t - M^2} \left(\eta^{\mu\nu} + \frac{p_1^\mu p_2^\nu}{M^2} \right) + (1 \leftrightarrow 2) \right] \right\}$$

6. Derive (14). [Hint: The procedure is exactly the same as that used to obtain (III.4.9).]
Write $\mathcal{L} = \frac{1}{2} A_\mu Q^{\mu\nu} A_\nu$ with $Q^{\mu\nu} = (\partial^2 + M^2)g^{\mu\nu} - [1 - (1/\xi)]\partial^\mu \partial^\nu$ or in momentum space
$Q^{\mu\nu} = -(k^2 - M^2)g^{\mu\nu} + [1 - (1/\xi)]k^\mu k^\nu$. The propagator is the inverse of $Q^{\mu\nu}$.

$$(14) \qquad \frac{-i}{k^2 - M^2 + i\varepsilon} \left[g_{\mu\nu} - (1 - \xi)\frac{k_\mu k_\nu}{k^2 - \xi M^2 + i\varepsilon} \right]$$

Solution:

The hint gave us the Fourier transform, so all we have to do is to find the matrix inverse
of $Q^{\mu\nu}$ in the Lorentz-index space. That is, solve $Q^{\mu\nu}(Q^{-1})_{\nu\rho} = \delta_\rho^\mu$ for $(Q^{-1})_{\nu\rho}$. Lorentz
invariance tells us that $(Q^{-1})_{\nu\rho} = a\, g_{\nu\rho} + b\, k_\nu k_\rho$, so plug this into the matrix inverse equation:

$$
\begin{aligned}
Q^{\mu\nu}(Q^{-1})_{\nu\rho} &= \left[(-k^2 + M^2)g^{\mu\nu} + (1 - 1/\xi)k^\mu k^\nu \right] \left[a\, g_{\nu\rho} + b\, k_\nu k_\rho \right] \\
&= (-k^2 + M^2)a\delta_\rho^\mu + (-k^2 + M^2)bk^\mu k_\rho + (1 - 1/\xi)ak^\mu k_\rho + (1 - 1/\xi)bk^2 k^\mu k_\rho \\
&= (-k^2 + M^2)a\delta_\rho^\mu + \left\{ (1 - 1/\xi)a + \left[-k^2 + M^2 + (1 - 1/\xi)k^2 \right] b \right\} k^\mu k_\rho \equiv \delta_\rho^\mu .
\end{aligned}
$$

The term multiplying δ_ρ^μ must equal 1, and the term multiplying $k^\mu k_\rho$ must equal zero. This
tells us that $a = 1/(-k^2 + M^2)$ and

$$
\begin{aligned}
(1 - 1/\xi)a + [-k^2 + M^2 + (1 - 1/\xi)k^2]b &= 0 \\
[k^2 - M^2 - (1 - 1/\xi)k^2]b &= +(1 - 1/\xi)a \\
[-M^2 + (1/\xi)k^2]b &= +(1 - 1/\xi)a \\
[k^2 - \xi M^2]b &= +(\xi - 1)a \\
b = (\xi - 1)\frac{1}{k^2 - \xi M^2}a &= (\xi - 1)\frac{1}{k^2 - \xi M^2}\left(\frac{-1}{k^2 - M^2} \right) .
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
(Q^{-1})_{\nu\rho} &= \left(\frac{-1}{k^2 - M^2} \right) g_{\nu\rho} + (\xi - 1)\frac{1}{k^2 - \xi M^2}\left(\frac{-1}{k^2 - M^2} \right) k_\nu k_\rho \\
&= \frac{-1}{k^2 - M^2} \left[g_{\nu\rho} + (\xi - 1)\frac{k_\nu k_\rho}{k^2 - \xi M^2} \right] .
\end{aligned}
$$

129

7. Work out the (...) in (13) and the Feynman rules for the various interaction vertices.

(13) $\qquad \mathcal{L} = \dfrac{\mu^4}{4\lambda} - \dfrac{1}{4}F_{\mu\nu}F^{\mu\nu} + \dfrac{1}{2}M^2 A_\mu A^\mu - MA_\mu\partial^\mu\varphi_2 + \dfrac{1}{2}[(\partial\varphi_1')^2 - 2\mu^2(\varphi_1')^2] + \dfrac{1}{2}(\partial\varphi_2)^2 + \ldots$

Solution:

We start with the $U(1)$-invariant scalar QED Lagrangian in Cartesian notation $\varphi = \frac{1}{\sqrt{2}}(\varphi_1 + i\varphi_2)$ as given in equation (12) of the text, on page 240:

$$\mathcal{L} = -\frac{1}{4}F^2 + \frac{1}{2}[(\partial\varphi_1 + eA\varphi_2)^2 + (\partial\varphi_2 - eA\varphi_1)^2] + \frac{\mu^2}{2}(\varphi_1^2 + \varphi_2^2) - \frac{\lambda}{4}(\varphi_1^2 + \varphi_2^2)^2$$

Now it's a matter of algebra:

$$(\partial\varphi_2 - eA\varphi_1)^2 = (\partial\varphi_2)^2 - 2eA\varphi_1\partial\varphi_2 + e^2A^2\varphi_1^2$$

$$(\partial\varphi_1 + eA\varphi_2)^2 = (\partial\varphi_1)^2 + 2eA\varphi_2\partial\varphi_1 + e^2A^2\varphi_2^2$$

So before spontaneous symmetry breaking, we have

$$\mathcal{L} = -\frac{1}{4}F^2 + \sum_{i=1}^{2}\left(\frac{1}{2}(\partial\varphi_i)^2 + \frac{1}{2}\mu^2\varphi_i^2 + \frac{1}{2}e^2A^2\varphi_i^2\right) - \frac{\lambda}{4}\left(\sum_{i=1}^{2}\varphi_i^2\right)^2 + eA_\mu(\varphi_2\partial^\mu\varphi_1 - \varphi_1\partial^\mu\varphi_2)$$

Now break the symmetry by writing $\varphi_1 = v + h$ ($h \equiv \varphi_1'$) and performing some more algebra:

$$\varphi_1^2 + \varphi_2^2 = (v + h)^2 + \varphi_2^2 = v^2 + 2vh + h^2 + \varphi_2^2$$

$$\begin{aligned}
(\varphi_1^2 + \varphi_2^2)^2 &= (v^2 + 2vh + h^2)^2 + 2(v^2 + 2vh + h^2)\varphi_2^2 + \varphi_2^4 \\
&= v^4 + 2v^2(2vh + h^2) + (2vh + h^2)^2 + 2v^2 + 4vh\varphi_2^2 + 2h^2\varphi_2^2 + \varphi_2^4 \\
&= v^4 + 4v^3h + 6v^2h^2 + 4vh^3 + h^4 + 2v^2 + 4vh\varphi_2^2 + 2h^2\varphi_2^2 + \varphi_2^4
\end{aligned}$$

Also, $\varphi_1\partial^\mu\varphi_2 = v\partial^\mu\varphi_2 + h\partial^\mu\varphi_2$. The Lagrangian is:

$$\begin{aligned}
\mathcal{L} = &-\frac{1}{4}F^2 + \frac{1}{2}(\partial h)^2 + \frac{1}{2}(\partial\varphi_2)^2 + \frac{\mu^2 + e^2A^2}{2}(v^2 + 2vh + h^2 + \varphi_2^2) \\
&- \frac{\lambda}{4}\left(v^4 + 4v^3h + 6v^2h^2 + 4vh^3 + h^4 + 2v^2\varphi_2^2 + 4vh\varphi_2^2 + 2h^2\varphi_2^2 + \varphi_2^4\right) \\
&+ eA_\mu(\varphi_2\partial^\mu h - h\partial^\mu\varphi_2) - evA_\mu\partial^\mu\varphi_2
\end{aligned}$$

Minimizing the potential with respect to v at the point for which all the fields are zero gives

$$\frac{\partial\mathcal{L}}{\partial v} = \mu^2 v - \lambda v^3 = 0 \implies \mu^2 = \lambda v^2$$

130

Therefore, $\frac{\mu^2}{2}2vh - \frac{\lambda}{4}4v^3h = 0$, which kills the terms linear in h, and $\frac{\mu^2}{2}h^2 - \frac{\lambda}{4}6v^2h^2 = \frac{1}{2}(1-3)\lambda v^2 h^2 = -\frac{1}{2}(\sqrt{2\lambda}\,v)^2 h^2$, so the mass of the physical h particle is $m_h = \sqrt{2\lambda}\,v$. The Lagrangian is:

$$\mathcal{L} = -\frac{1}{4}F^2 + \frac{1}{2}(ev)^2 A^2 + \frac{1}{2}(\partial h)^2 - \frac{1}{2}m_h^2 h^2 - \frac{\lambda}{4}(4vh^3 + h^4 + 4vh\varphi_2^2 + 2h^2\varphi_2^2 + \varphi_2^4)$$
$$+ \frac{1}{2}(\partial\varphi_2)^2 - evA\partial\varphi_2 + eA(\varphi_2\partial h - h\partial\varphi_2) + \frac{1}{2}e^2 A^2(2vh + h^2 + \varphi_2^2)$$

As explained on p. 240, we add the gauge fixing term $\mathcal{L}_{\text{gf}} = -\frac{1}{2\xi}(\partial A + \xi M\varphi_2)^2$, where $M = ev$ is the mass of the photon after spontaneous symmetry breaking. Since

$$\mathcal{L}_{\text{gf}} = -\frac{1}{2\xi}[(\partial A)^2 + 2\xi M\partial A\varphi_2 + \xi^2 M^2\varphi_2^2]$$
$$= -\frac{1}{2\xi}(\partial A)^2 - M\partial A\varphi_2 - \frac{1}{2}\xi M^2\varphi_2^2$$
$$= -\frac{1}{2\xi}(\partial A)^2 + MA\partial\varphi_2 - \frac{1}{2}\xi M^2\varphi_2^2 + \text{(total derivative)}$$

the $A\partial\varphi_2$ term cancels from the Lagrangian (by design), and we get:

$$\mathcal{L} = -\frac{1}{4}F^2 - \frac{1}{2\xi}(\partial A)^2 + \frac{1}{2}M^2 A^2 \quad \leftarrow \quad \text{(yields massive vector boson propagator in } R_\xi \text{ gauge)}$$

$$+ \frac{1}{2}(\partial h)^2 - \frac{1}{2}m_h^2 h^2 \quad \leftarrow \quad \text{(yields scalar propagator for } h\text{)}$$

$$+ \frac{1}{2}(\partial\varphi_2)^2 - \frac{1}{2}\xi M^2\varphi_2^2 \quad \leftarrow \quad \text{(yields scalar propagator for Goldstone boson } \varphi_2\text{)}$$

$$- \lambda v\, h^3 - \frac{1}{4}\lambda\, h^4 - \frac{1}{4}\lambda\varphi_2^4 \quad \leftarrow \quad \text{(cubic and quartic self-interactions for } h, \text{ quartic for } \varphi_2\text{)}$$

$$+ eM\, hA_\mu A^\mu + \frac{1}{2}e^2\,(h^2 + \varphi_2^2)A_\mu A^\mu \quad \leftarrow \quad \text{(photon-scalar interactions)}$$

$$- \lambda v\, h\varphi_2^2 - \frac{1}{2}\lambda\, h^2\varphi_2^2 \quad \leftarrow \quad (h\text{-}\varphi_2 \text{ interactions)}$$

$$+ e(\varphi_2\partial_\mu h - h\partial_\mu\varphi_2)A^\mu \quad \leftarrow \quad \text{(photon coupling to scalar electromagnetic current)}$$

Using solid lines for h, dashed lines for φ_2 and wavy lines for A^μ, the interaction vertices in momentum space are:

$$= -i3!\lambda v$$

$$= -i3!\lambda$$

$$= -i3!\lambda$$

$$= +i(2e^2)\eta^{\mu\nu}$$

$$= +i(2e^2)\eta^{\mu\nu}$$

$$= -i2\lambda v$$

$$= +i2eM\eta^{\mu\nu}$$

$$= -i2\lambda$$

$$\mu = +e(k_1 - k_2)^\mu$$
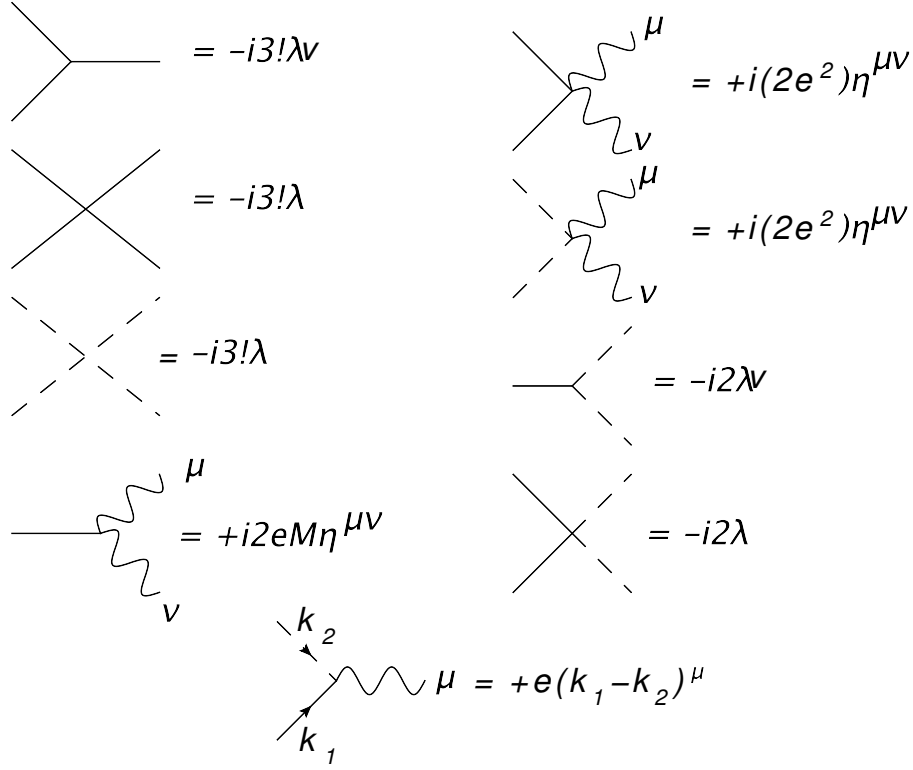
where in the last diagram, the arrows indicate that both momenta flow into the vertex.

8. Using the Feynman rules derived in exercise IV.6.7 calculate the amplitude for the physical process $\varphi_1' + \varphi_1' \to A + A$ and show that the dependence on ξ cancels out. Compare with the result in exercise IV.6.5. [Hint: There are two diagrams, one with A exchange and the other with φ_2 exchange.]

Solution:

As for IV.6.5, the amplitude is a sum of diagrams: $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2 + \mathcal{M}_5 + \mathcal{M}_5'$, where $\mathcal{M}_1$ and $\mathcal{M}_2$ are identical to those of IV.6.5 because the vertices are identical. The next two diagrams are the ones mentioned in the hint, one with t-channel φ_2 exchange and one with t-channel A_μ exchange; we define both of these as contained within $\mathcal{M}_5$. Finally, we define $\mathcal{M}_5'$ as the crossed version of $\mathcal{M}_5$. The four diagrams contained in $\mathcal{M}_5 + \mathcal{M}_5'$ must conspire to give the same result as the sum $\mathcal{M}_3 + \mathcal{M}_4$ did in IV.6.5, since the two scattering processes should be identical.

The gauge boson propagator in R_ξ gauge is

$$i\Delta_{\mu\nu}(q) \equiv \text{wavy line} = \frac{-i}{q^2 - M^2 + i\varepsilon}\left[\eta_{\mu\nu} - (\xi - 1)\frac{q_\mu q_\nu}{q^2 - \xi M^2 + i\varepsilon}\right]$$

132

and the φ_2 propagator is

$$i\Delta_{\mu\nu}(q) \equiv \text{dashed line} = \frac{i}{q^2 - \xi M^2 + i\varepsilon} \, .$$

The partial amplitude $\mathcal{M}_5$ is (again $q = p_1 - p_{1'} = p_{2'} - p_2$):

$$
\begin{aligned}
\mathcal{M}_5 &= \varphi_2\text{-exchange} + A_\mu\text{-exchange} \\
&= \varepsilon_{1'\mu}\varepsilon_{2'\nu} \left[e(p_1 + q)^\mu i\Delta(q)\, e(p_2 - q)^\nu + (i2eM\eta^{\mu\rho})i\Delta_{\rho\sigma}(q)(i2eM\eta^{\sigma\nu}) \right] \\
&= ie^2\varepsilon_{1'\mu}\varepsilon_{2'\nu} \left[\frac{(2p_1 - p_{1'})^\mu(2p_2 - p_{2'})^\nu}{q^2 - \xi M^2} + (2M)^2\frac{1}{q^2 - M^2}\left(\eta^{\mu\nu} + (\xi - 1)\frac{q^\mu q^\nu}{q^2 - \xi M^2} \right) \right] \\
&= 4ie^2\varepsilon_{1'\mu}\varepsilon_{2'\nu} \left[\frac{M^2}{t - M^2}\eta^{\mu\nu} + \left(1 - (\xi - 1)\frac{M^2}{t - M^2} \right)\frac{p_1^\mu p_2^\mu}{t - \xi M^2} \right]
\end{aligned}
$$

where $t \equiv q^2 = (p_1 - p_{1'})^2 = (p_{2'} - p_2)^2$, and again we have used gauge invariance in the form of $\varepsilon_{1'} \cdot p_{1'} = 0$ and $\varepsilon_{2'} \cdot p_{2'} = 0$ to write $\varepsilon_{1'} \cdot q = \varepsilon_{1'} \cdot p_1$ and $\varepsilon_{2'} \cdot q = -\varepsilon_{2'} \cdot p_2$. Now simplify the quantity in parentheses:

$$1 - (\xi - 1)\frac{M^2}{t - M^2} = \frac{t - M^2 - (\xi - 1)M^2}{t - M^2} = \frac{t - \xi M^2}{t - M^2} \, .$$

The ξ-dependent numerator cancels the $t - \xi M^2$ in the denominator of $p_1^\mu p_2^\nu$, so we find

$$\mathcal{M}_5 = i4e^2\varepsilon_{1'\mu}\varepsilon_{2'\nu}\frac{M^2}{t - M^2}\left(\eta^{\mu\nu} + \frac{p_1^\mu p_2^\mu}{M^2} \right) \, .$$

The crossed diagrams can be obtained by switching $p_1 \leftrightarrow p_2$ to get

$$\mathcal{M}_5' = i4e^2\varepsilon_{1'\mu}\varepsilon_{2'\nu}\frac{M^2}{u - M^2}\left(\eta^{\mu\nu} + \frac{p_2^\mu p_1^\mu}{M^2} \right)$$

where $u \equiv (p_{2'} - p_1)^2 = (p_2 - p_{1'})^2$.

Thus we find $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2 + \mathcal{M}_3 + \mathcal{M}_4 = \mathcal{M}_1 + \mathcal{M}_2 + \mathcal{M}_5 + \mathcal{M}_5'$, as we must since the two theories are equivalent.

9. Consider the theory defined in (12) with $\mu = 0$. Using the result of exercise IV.3.5 show that

$$V_{\text{eff}}(\varphi) = \frac{1}{4}\lambda\varphi^4 + \frac{1}{64\pi^2}(10\lambda^2 + 3e^4)\varphi^4\left(\ln\frac{\varphi^2}{M^2} - \frac{25}{6}\right) + \ldots$$

where $\varphi^2 = \varphi_1^2 + \varphi_2^2$. This potential has a minimum away from $\varphi = 0$ and thus the gauge symmetry is spontaneously broken by quantum fluctuations. In chapter IV.3 we did not have the e^4 term and argued that the minimum we got there was not to be trusted. But here we can balance the $\lambda\varphi^4$ against $e^4\varphi^4\ln(\varphi^2/M^2)$ for λ of the same order of magnitude as e^4. The minimum can be trusted. Show that the spectrum of this theory consists of a massive scalar boson and a massive vector boson, with

$$\frac{m_{\text{scalar}}^2}{m_{\text{vector}}^2} = \frac{3}{2\pi}\frac{e^2}{4\pi}.$$

For help, see S. Coleman and E. Weinberg, *Phys. Rev.* D7: 1888, 1973.

Solution:

The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}(\partial\varphi_1 - eA\varphi_2)^2 + \frac{1}{2}(\partial\varphi_2 + eA\varphi_1)^2 - \frac{1}{4!}\lambda(\varphi_1^2 + \varphi_2^2)^2 + \mathcal{L}_{\text{ct}}$$

where $\mathcal{L}_{\text{ct}}$ includes the counterterms. Since the result will depend only on $\varphi_c^2 = \varphi_{1c}^2 + \varphi_{2c}^2$, we need only to consider diagrams whose external lines are φ_1. The diagrams we need to consider are of the form
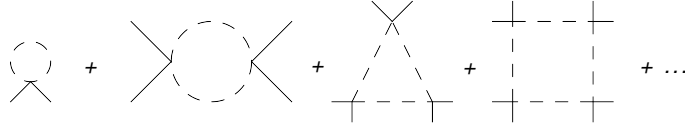
where the internal dotted lines are either φ_1, φ_2, or A_μ.

The text already calculated the scalar contributions to the effective potential. In problem IV.3.5, we showed that the photon contribution is the same as the scalar contribution, except with the replacement $\lambda \to 2e^2$ and an overall factor of 3. Thus the photon contribution to V_{eff} is

$$V_{\text{eff}}^{(A)}(\varphi_c) = 3 \times \frac{1}{256\pi^2}[2e^2]^2\varphi_c^4\left(\ln\frac{\varphi_c^2}{M^2} - \frac{25}{6}\right)$$

$$= \frac{3e^4}{64\pi^2}\varphi_c^4\left(\ln\frac{\varphi_c^2}{M^2} - \frac{25}{6}\right).$$

Summing up all contributions, we find

$$V_{\text{eff}}(\varphi_c) = \frac{1}{4!}\lambda\varphi_c^4 + \left(\frac{5\lambda^2}{1152\pi^2} + \frac{3e^4}{64\pi^2}\right)\varphi_c^4\left(\ln\frac{\varphi_c^2}{M^2} - \frac{25}{6}\right)$$

134

where the couplings are renormalized at the scale M.

Now treat λ as a coupling of order e^4, so that the term of order $\lambda^2 = O(e^8)$ should be dropped. Then we have

$$V_{\text{eff}}(\varphi_c) = \frac{1}{4!}\lambda\varphi_c^4 + \frac{3e^4}{64\pi^2}\varphi_c^4\left(\ln\frac{\varphi_c^2}{M^2} - \frac{25}{6}\right)$$

where the couplings are renormalized at the scale M.

Let the value $\varphi_c = v$ be the location of the minimum of $V_{\text{eff}}(\varphi_c)$, or in other words $V'_{\text{eff}}(v) = 0$. Furthermore, let us choose $M = v$. Then the equation $V'_{\text{eff}}(v) = 0$ implies

$$\lambda = \frac{33e^4}{8\pi^2} .$$

Putting this back into the effective potential results in

$$V_{\text{eff}}(\varphi_c) = \frac{3e^4}{64\pi^2}\varphi_c^4\left(\ln\frac{\varphi_c^2}{v^2} - \frac{1}{2}\right) .$$

The squared mass of the scalar is given by

$$m_S^2 \equiv V''_{\text{eff}}(v) = \frac{3e^4}{8\pi^2}v^2 .$$

The mass of the vector after spontaneous symmetry breaking can be read off from the Lagrangian as

$$m_V = ev .$$

Dividing these, we arrive at the relation

$$\frac{m_S^2}{m_V^2} = \frac{3\alpha}{2\pi}$$

where $\alpha \equiv e^2/(4\pi)$.

IV.7 Chiral Anomaly

1. Derive (11) from (9). The momentum factors $k_{1\lambda}$ and $k_{2\sigma}$ in (9) become the two derivatives in $F_{\mu\nu}F_{\lambda\sigma}$ in (11).

$$q_\lambda \Delta^{\lambda\mu\nu}(a, k_1, k_2) = \frac{i}{2\pi^2}\varepsilon^{\mu\nu\lambda\sigma} k_{1\lambda} k_{2\sigma} \qquad (9)$$

$$\partial_\mu J_5^\mu = \frac{e^2}{(4\pi)^2}\varepsilon^{\mu\nu\lambda\sigma} F_{\mu\nu}F_{\lambda\sigma} \qquad (11)$$

Solution:

We will derive this relation by comparing the matrix elements of both operators in the canonical formalism, and finding that they are the same up to numerical factors. We are interested in processes for which an operator causes two photons to pop out of the vacuum, or in other words we want matrix elements of the form

$$\langle k, a;\ k', b|\mathcal{O}(0)|0\rangle$$

where a labels the polarization of the outgoing photon with momentum k, and b labels the polarization of the outgoing photon with momentum k'. By translation invariance, $\langle k, a;\ k', b|\mathcal{O}(x)|0\rangle = e^{-i(k+k')x}\langle k, a;\ k', b|\mathcal{O}(0)|0\rangle$, so it is sufficient to consider the operator at the origin $x = 0$. The first matrix element we will consider is for the operator $\mathcal{O}(x) = -i\partial_\mu J_5^\mu(x)$, for which all of the work has already been done in the chapter:

$$\langle k, a;\ k', b|(-i)\partial_\mu J_5^\mu(0)|0\rangle = \varepsilon_\mu^a(k)^* \varepsilon_\nu^b(k')^* e^2 (-i) q_\lambda \Delta^{\lambda\mu\nu}(a, k, k') = \varepsilon_\mu^a(k)^* \varepsilon_\nu^b(k')^* \frac{e^2}{2\pi^2}\varepsilon^{\mu\nu\lambda\sigma} k_\lambda k'_\sigma\ .$$

The factor of $-i$ came from the Fourier replacement $\partial_\mu \to iq_\mu$. Now consider the matrix element for $\mathcal{O}(x) = \varepsilon^{\mu\nu\rho\sigma}\partial_\mu A_\nu(x)\partial_\rho A_\sigma(x)$. We will need the plane-wave expansion for the photon field:

$$A_\nu(x) = \sum_a \int \frac{d^3p}{(2\pi)^3 2\omega_p}\left[a_p \varepsilon_\nu^a(p)\, e^{-ipx} + a_p^\dagger \varepsilon_\nu^a(p)^*\, e^{+ipx}\right]\ .$$

The spacetime derivative at the origin is

$$\partial_\mu A_\nu(0) = \sum_a \int \frac{d^3p}{(2\pi)^3 2\omega_p}(-ip_\mu)\left[a_p \varepsilon_\nu^a(p) - a_p^\dagger \varepsilon_\nu^a(p)^*\right]\ .$$

We will need the following properties of the creation and annihilation operators acting on the vacuum state:

$$a_k|0\rangle = 0$$
$$a_k a_{k'}^\dagger|0\rangle = (2\pi)^3(2\omega_k)\delta^3(\vec{k} - \vec{k}')$$
$$\langle k, k'|0\rangle = \langle 0|a_{k'}a_k|0\rangle = 0$$
$$\langle k, a;\ k', b|p, c;\ p', d\rangle =$$
$$[(2\pi)^3 2\omega_k][(2\pi)^3 2\omega_{k'}]\left[\delta^{ac}\delta^{bd}\delta^3(\vec{k} - \vec{p})\delta^3(\vec{k}' - \vec{p}') + \delta^{ad}\delta^{bc}\delta^3(\vec{k} - \vec{p}')\delta^3(\vec{k}' - \vec{p})\right]$$

136

In what follows, define $(dp) \equiv \frac{d^3p}{(2\pi)^3 2\omega_p}$ and suppress the polarization labels on the outgoing states to save room. Now compute the matrix element:

$$\langle k, k' | \varepsilon^{\mu\nu\rho\sigma} \partial_\mu A_\nu(0) \partial_\rho A_\sigma(0) | 0 \rangle$$

$$= \varepsilon^{\mu\nu\rho\sigma} \sum_{c,d} \int (dp)(dp')(-ip_\mu)(-ip'_\rho)\langle k, k' | \left[a_p \varepsilon_\nu^c(p) - a_p^\dagger \varepsilon_\nu^c(p)^* \right] \left[a_{p'} \varepsilon_\sigma^d(p') - a_{p'}^\dagger \varepsilon_\sigma^d(p')^* \right] | 0 \rangle$$

$$= (-1)^2 \varepsilon^{\mu\nu\rho\sigma} \sum_{c,d} \int (dp)(dp') p_\mu p'_\rho \langle k, k' || \left[a_p \varepsilon_\nu^c(p) - a_p^\dagger \varepsilon_\nu^c(p)^* \right] a_{p'}^\dagger | 0 \rangle \varepsilon_\sigma^d(p')^*$$

$$= - \varepsilon^{\mu\nu\rho\sigma} \sum_{c,d} \int (dp)(dp') p_\mu p'_\rho \langle k, k' | p, p' \rangle \varepsilon_\nu^c(p)^* \varepsilon_\sigma^d(p')^*$$

$$= -2 \varepsilon^{\mu\nu\rho\sigma} k_\mu k'_\rho \varepsilon_\nu^a(k)^* \varepsilon_\sigma^b(k')^*$$

$$= -2 \varepsilon^{\rho\mu\sigma\nu} k_\rho k'_\sigma \varepsilon_\mu^a(k)^* \varepsilon_\nu^b(k')^*$$

$$= -2 (-1)^3 \varepsilon^{\mu\nu\rho\sigma} k_\rho k'_\sigma \varepsilon_\mu^a(k)^* \varepsilon_\nu^b(k')^*$$

$$= +2 \varepsilon^{\mu\nu\rho\sigma} \varepsilon_\mu^a(k)^* \varepsilon_\nu^b(k')^* k_\rho k'_\sigma$$

Recall the definition of the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Then

$$\varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} = 4\varepsilon^{\mu\nu\rho\sigma} \partial_\mu A_\nu \partial_\rho A_\sigma$$

so in total we have

$$\langle k, a; k', b | \varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu}(0) F_{\rho\sigma}(0) | 0 \rangle = +8 \varepsilon^{\mu\nu\rho\sigma} \varepsilon_\mu^a(k)^* \varepsilon_\nu^b(k')^* k_\rho k'_\sigma .$$

Therefore

$$\langle k, a; k', b | \partial_\mu J_5^\mu(0) | 0 \rangle = \frac{e^2}{2\pi^2} \varepsilon^{\mu\nu\lambda\sigma} \varepsilon_\mu^a(k)^* \varepsilon_\nu^b(k')^* k_\lambda k'_\sigma = \frac{e^2}{2\pi^2} \left(\frac{1}{8} \langle k, a; k', b | \varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu}(0) F_{\rho\sigma}(0) | 0 \rangle \right) .$$

We therefore have the operator equation

$$\partial_\mu J_5^\mu = \frac{e^2}{16\pi^2} \varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} .$$

2. Following the reasoning in chapter IV.2 and using the erroneous (10) show that the decay amplitude for the decay $\pi^0 \to \gamma + \gamma$ would vanish in the ideal world in which the π^0 is massless. Since the π^0 does decay and since our world is close to the ideal world, this provided the first indication historically that (10) cannot possibly be valid.

Solution:

The decay $\pi^0 \to \gamma\gamma$ has an amplitude

$$A(\pi^0 \to \gamma\gamma) \propto k_\mu \langle \gamma_1 \gamma_2 | J_5^\mu(0) | \pi^0(k) \rangle$$

where k^μ is the momentum of the decaying neutral pion. This amplitude has one initial hadron and no final hadrons. By translation invariance, we have

$$\langle \gamma_1 \gamma_2 | J_5^\mu(x) | \pi^0(k) \rangle = \langle \gamma_1 \gamma_2 | J_5^\mu(0) | \pi^0(k) \rangle e^{-ik \cdot x}$$

As in chapter IV.2, we have $\langle \gamma_1 \gamma_2 | J_5^\mu(0) | \pi^0(k) \rangle = f_0 k^\mu$ and so $k_\mu \langle \gamma_1 \gamma_2 | J_5^\mu(0) | \pi^0(k) \rangle = f_0 k^2 = f_0 m_{\pi^0}^2$.

Since $\langle \gamma_1 \gamma_2 | \partial_\mu J_5^\mu(x) | \pi^0(k) \rangle = -ik_\mu \langle \gamma_1 \gamma_2 | J_5^\mu(0) | \pi^0(k) \rangle e^{-ik \cdot x} = -if_0 m_{\pi^0}^2 e^{-ik \cdot x}$, we deduce (incorrectly) that $\partial_\mu J_5^\mu(x) = 0$ if and only if $f_0 m_{\pi^0}^2 = 0$.

In other words, $m_{\pi^0} \to 0$ would seem to imply $k_\mu \langle \gamma_1 \gamma_2 | J_5^\mu(0) | \pi^0(k) \rangle \to 0$. Since $A(\pi^0 \to \gamma\gamma)$ is proportional to that, we deduce incorrectly that the decay $\pi^0 \to \gamma\gamma$ cannot occur.

3. Repeat all the calculations in the text for the theory $\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi$.

Solution (due to J. Feinberg):

The momentum space amplitude for massive fermions is

$$\Delta^{\lambda\mu\nu}(k_1, k_2) = i \int \frac{d^4 p}{(2\pi)^4} \mathrm{tr}\left[\gamma^\lambda \gamma^5 \frac{1}{\not{p} - \not{q} - m} \gamma^\nu \frac{1}{\not{p} - \not{k}_1 - m} \gamma^\mu \frac{1}{\not{p} - m} + (\mu, k_1 \leftrightarrow \nu, k_2) \right],$$

which is linearly divergent by power counting. Therefore, we have to specify how we will regulate and evaluate it. Following the text, we will set up all divergent integrals as integrals of total derivatives, using the formula

$$\int d^4 p [f(p + a) - f(p)] = \lim_{P \to \infty} i\, a_\mu P^\mu (2\pi^2 P^2) f(P)$$

evaluated symmetrically over the 3-sphere at infinity (so that $P^\mu P^\nu / P^2 \to \frac{1}{4} g^{\mu\nu}$, etc.). Furthermore, we will exchange $(\mu, k_1 \leftrightarrow \nu, k_2)$ at the last step in any calculation, and we define the momentum space amplitude as the amplitude given above with loop momentum p shifted by a chosen so that the vector currents are conserved: $k_{1\mu} \Delta^{\lambda\mu\nu} = 0$ and $k_{2\nu} \Delta^{\lambda\mu\nu} = 0$. Technically, also, when $P \to \infty$, we should be taking that limit in Euclidean space, so we don't get complications from null vectors, and we can always take $P^2 \gg m^2$.

So let us find the difference of the shifted and unshifted amplitudes, as in the text:

$$\Delta^{\lambda\mu\nu}(a) - \Delta^{\lambda\mu\nu}(0) = \frac{i}{(2\pi)^4}(i2\pi^2) a^\rho \lim_{P \to \infty} P_\rho P^2 \frac{N}{D}$$

where the numerator is

$$N = \mathrm{tr}\left[\gamma^\lambda \gamma^5 (\not{P} - \not{q} + m) \gamma^\nu (\not{P} - \not{k}_1 + m) \gamma^\mu (\not{P} + m) \right]$$

138

and the denominator is

$$D = [(P-q)^2 - m^2][(P-k_1)^2 - m^2][P^2 - m^2] \, .$$

To this we also add $(\mu, k_1 \leftrightarrow \nu, k_2)$ as in the text. Only the terms in the numerator with six powers of P survive, so we need $\mathrm{tr}[\gamma^\lambda \gamma^5 \slashed{P} \gamma^\nu \slashed{P} \gamma^\mu \slashed{P}] = 2P^\mu \mathrm{tr}[\gamma^\lambda \gamma^5 \slashed{P} \gamma^\nu \slashed{P}] - P^2 \mathrm{tr}[\gamma^\lambda \gamma^5 \slashed{P} \gamma^\nu \gamma^\mu]$ using the gamma-matrix anticommutator. Then $\mathrm{tr}[\gamma^5 \gamma^\sigma \gamma^\nu \gamma^\mu \gamma^\lambda] = 4i\varepsilon^{\sigma\nu\mu\lambda}$, so the first term vanishes. Then

$$\Delta^{\lambda\mu\nu}(a) - \Delta^{\lambda\mu\nu}(0) = \frac{4i}{8\pi^2} a^\rho \varepsilon^{\sigma\nu\mu\lambda} \lim_{P\to\infty} P_\rho P_\sigma + (\mu, k_1 \leftrightarrow \nu, k_2)$$

$$= \frac{i}{8\pi^2} \varepsilon^{\sigma\nu\mu\lambda} a_\sigma + (\mu, k_1 \leftrightarrow \nu, k_2) \, .$$

Now we need to find the shift variable a so that the vector currents are conserved. The only independent momenta available in the problem are k_1 and k_2, and Bose symmetry implies that we take $a = \beta(k_1 - k_2)$. Thus

$$\Delta^{\lambda\mu\nu}(a) - \Delta^{\lambda\mu\nu}(0) = \frac{i\beta}{4\pi^2} \varepsilon^{\lambda\mu\nu\sigma}(k_1 - k_2)_\sigma$$

as in the text. Now we can enforce vector current conservation. Using $\slashed{k}_1 = (\slashed{p} - m) - (\slashed{p} - \slashed{k}_1 - m) = (\slashed{p} - \slashed{k}_2 - m) - (\slashed{p} - \slashed{q} - m)$, we can write

$$k_{1\mu}\Delta^{\lambda\mu\nu}(0) = i \int \frac{d^4p}{(2\pi)^4} \mathrm{tr}\left[\gamma^\lambda \gamma^5 \frac{1}{\slashed{p} - \slashed{q} - m} \gamma^\nu \frac{1}{\slashed{p} - \slashed{k}_1 - m} - \gamma^\lambda \gamma^5 \frac{1}{\slashed{p} - \slashed{q} - m} \gamma^\nu \frac{1}{\slashed{p} - m}\right]$$

$$+ i \int \frac{d^4p}{(2\pi)^4} \mathrm{tr}\left[\gamma^\lambda \gamma^5 \frac{1}{\slashed{p} - \slashed{q} - m} \gamma^\nu \frac{1}{\slashed{p} - m} - \gamma^\lambda \gamma^5 \frac{1}{\slashed{p} - \slashed{k}_2 - m} \gamma^\nu \frac{1}{\slashed{p} - m}\right]$$

$$= -ik_1^\rho \int \frac{d^4p}{(2\pi)^4} \frac{\partial}{\partial p^\rho} \mathrm{tr}\left[\gamma^\lambda \gamma^5 \frac{1}{\slashed{p} - \slashed{k}_2 - m} \gamma^\nu \frac{1}{\slashed{p} - m}\right]$$

$$= 4k_1^\rho \varepsilon^{\sigma\nu\tau\lambda} \int \frac{d^4p}{(2\pi)^4} \frac{\partial}{\partial p^\rho} \frac{(p - k_2)_\sigma p_\tau}{[(p - k_2)^2 - m^2][p^2 - m^2]}$$

$$= -\frac{4i}{8\pi^2} k_1^\rho k_{2\sigma} \varepsilon^{\lambda\nu\sigma\tau} \lim_{P\to\infty} \frac{P_\rho P_\tau}{P^2}$$

$$= +\frac{i}{8\pi^2} \varepsilon^{\lambda\nu\tau\sigma} k_{1\tau} k_{2\sigma} \, .$$

In the third equality, we have used the fact that traces of γ^5 with fewer than four other gamma matrices vanish. Putting this all together, we obtain:

$$k_{1\mu}\Delta^{\lambda\mu\nu}(a) = k_{1\mu}\Delta^{\lambda\mu\nu}(0) + \frac{i\beta}{4\pi^2} \varepsilon^{\lambda\mu\nu\sigma} k_{1\mu} k_{2\sigma}$$

$$= \frac{i}{8\pi^2}(1 + 2\beta) \varepsilon^{\lambda\mu\nu\sigma} k_{1\mu} k_{2\sigma}$$

139

Demanding vector current conservation means setting this quantity to be zero, so as in the text we need to set $\beta = -1/2$.

Finally, we can calculate the divergence of the axial current. We use

$$\slashed{q} = (\slashed{p} - m) - (\slashed{p} - \slashed{q} + m) + 2m$$

where the difference in sign of m in the second term is due to anticommuting the slashed momenta with γ^5. After some algebra,

$$q_\lambda \Delta^{\lambda\mu\nu} = -ik_1^\rho \int \frac{d^4p}{(2\pi)^4} \frac{\partial}{\partial p^\rho} \text{tr} \left[\gamma^5 \frac{1}{\slashed{p} - \slashed{k}_2 - m} \gamma^\nu \frac{1}{\slashed{p} - m} \gamma^\mu \right] + (\mu, k_1 \leftrightarrow \nu, k_2) + 2m\Delta^{\mu\nu}$$

where

$$\Delta^{\mu\nu}(k_1, k_2, m) = i \int \frac{d^4p}{(2\pi)^4} \text{tr} \left[\gamma^5 \frac{1}{\slashed{p} - \slashed{q} - m} \gamma^\nu \frac{1}{\slashed{p} - \slashed{k}_1 - m} \gamma^\mu \frac{1}{\slashed{p} - m} \right] .$$

The first two terms evaluate just as in the book to give

$$\frac{4i}{8\pi^2} k_1^\rho (-k_{2\sigma}) \varepsilon^{\sigma\nu\tau\mu} \lim_{P\to\infty} \frac{P_\sigma P_\tau}{P^2} + (\mu, k_1 \leftrightarrow \nu, k_2) = \frac{i}{4\pi^2} \varepsilon^{\mu\nu\sigma\tau} k_{1\sigma} k_{2\tau} .$$

Thus, our properly regulated amplitude is

$$q_\lambda \Delta^{\lambda\mu\nu}(a) = \frac{i}{2\pi^2} \varepsilon^{\mu\nu\sigma\tau} k_{1\sigma} k_{2\tau} + 2m\Delta^{\mu\nu}(k_1, k_2, m) .$$

Now we just need to calculate $\Delta^{\mu\nu}(k_1, k_2, m)$. We'll work on the first integral and then perform the Bose swap. We have

$$i \int \frac{d^4p}{(2\pi)^4} \frac{\text{tr}[\gamma^5(\slashed{p} - \slashed{q} + m)\gamma^\nu(\slashed{p} - \slashed{k}_1 + m)\gamma^\mu(\slashed{p} + m)]}{[(p-q)^2 - m^2][(p-k_1)^2 - m^2][p^2 - m^2]} .$$

From the trace, the only terms that don't vanish are linear in m, or

$$m \, \text{tr} \left[\gamma^5 \gamma^\nu (\slashed{p} - \slashed{k}_1) \gamma^\mu \slashed{p} + \gamma^5 (\slashed{p} - \slashed{q}) \gamma^\nu \gamma^\mu \slashed{p} + \gamma^5 (\slashed{p} - \slashed{q}) \gamma^\nu (\slashed{p} - \slashed{k}_1) \gamma^\mu \right] ,$$

which become after some algebra and a bunch of cancellations

$$-4i \, m \, \varepsilon^{\mu\nu\sigma\tau} k_{1\sigma} k_{2\tau} .$$

Thus

$$\Delta^{\mu\nu}(k_1, k_2, m) = 4m \, \varepsilon^{\mu\nu\sigma\tau} k_{1\sigma} k_{2\tau} [I(q, k_1, m) + I(q, k_2, m)]$$

where

$$I(q, k, m) = \int \frac{d^4p}{(2\pi)^4} \frac{1}{[(p-q)^2 - m^2][(p-k)^2 - m^2][p^2 - m^2]}$$

is superficially convergent, so we can just evaluate it and shift variables as desired. Using Feynman parameters and shifting variables, we get

$$I(q, k, m) = 2 \int_0^1 dx\, dy\, dz \int \frac{d^4p}{(2\pi)^4} \frac{\delta(x+y+z-1)}{[p^2 - m^2 - 2p \cdot qx - 2p \cdot ky + q^2x + k^2y]^3}$$
$$= 2 \int_0^1 dx\, dy\, dz \int \frac{d^4\ell}{(2\pi)^4} \frac{\delta(x+y+z-1)}{[\ell^2 - m^2 + q^2x(1-x) + k^2y(1-y) + 2q \cdot kxy]^3}$$

for $\ell \equiv p - qx - ky$. Integrating gives

$$I(q, k, m) = -\frac{i}{16\pi^2} \int_0^1 dx\, dy\, dz\, \frac{\delta(x+y+z-1)}{m^2 - q^2x(1-x) - k^2y(1-y) - 2q \cdot kxy} \, .$$

It is relatively easy to put the expression back together again. The extra term is just what we would expect from $\langle \bar{\psi}\gamma^5\psi\, J^\mu J^\nu \rangle$.

6. Discuss the anomaly by studying the amplitude

$$\langle 0|T[J_5^\lambda(0) J_5^\mu(x_1) J_5^\nu(x_2)]|0\rangle$$

given in lowest orders by triangle diagrams with axial currents at each vertex. [Hint: Call the momentum space amplitude $\Delta_5^{\lambda\mu\nu}(k_1, k_2)$.] Show by using $(\gamma^5)^2 = 1$ and Bose symmetry that

$$\Delta_5^{\lambda\mu\nu}(k_1, k_2) = \frac{1}{3}[\Delta^{\lambda\mu\nu}(a, k_1, k_2) + \Delta^{\mu\nu\lambda}(a, k_2, -q) + \Delta^{\nu\lambda\mu}(a, -q, k_1)].$$

Now use (9) to evaluate $q_\lambda \Delta_5^{\lambda\mu\nu}(k_1, k_2)$.

Solution (due to J. Feinberg):

The momentum space amplitude, by analogy with the one studied in the text, is

$$\Delta_5^{\lambda\mu\nu} = i \int \frac{d^4p}{(2\pi)^4} \text{tr}\left[\gamma^\lambda\gamma^5 \frac{1}{\not{p}-\not{q}} \gamma^\nu\gamma^5 \frac{1}{\not{p}-\not{k}_1} \gamma^\mu\gamma^5 \frac{1}{\not{p}}\right] + (\mu, k_1 \leftrightarrow \nu, k_2) \, .$$

We have only the two terms because there are only two circular permutations of the 3 vertices. This is of course a divergent integral, so we must specify how we will regulate it. First off, by Bose symmetry, since the 3 currents are identical, we must choose a regulator which is invariant under $(\lambda, -q \leftrightarrow \mu, k_1)$ and $(\lambda, -q \leftrightarrow \nu, k_2)$ as well as $(\mu, k_1 \leftrightarrow \nu, k_2)$. (The sign on q is because q is incoming as opposed to outgoing.) A simple way to implement this type of regularization is to define $\Delta_5^{\lambda\mu\nu}(a)$ as the expression $\Delta_5^{\lambda\mu\nu}$ above but with the integration variable p shifted to $p + a$ and then define our regulated amplitude as

$$\Delta_5^{\lambda\mu\nu} = \frac{1}{3}[\Delta_5^{\lambda\mu\nu}(0) + \Delta_5^{\mu\nu\lambda}(k_1) + \Delta_5^{\nu\lambda\mu}(q)]$$

141

with the permutation of indices indicating that we have also used the cyclic property of the trace. Using the second or third terms would be equivalent to evaluating the Feynman diagram starting at a different vertex.

From looking at the integral expression for $\Delta_5^{\lambda\mu\nu}$ above, we can anticommute the second γ^5 matrix through $1/(\not{p}-\not{k}_1)$ and $1/(\not{p}-\not{k}_2)$ and then γ^μ and γ^ν, and use $(\gamma^5)^2 = 1$ to see that

$$\Delta_5^{\lambda\mu\nu}(0) = \Delta^{\lambda\mu\nu}(k_1, k_2) \ , \ \Delta_5^{\mu\nu\lambda}(k_1) = \Delta^{\mu\nu\lambda}(k_2, -q) \ , \ \Delta_5^{\nu\lambda\mu}(q) = \Delta^{\nu\lambda\mu}(-q, k_1) \ .$$

We should now use the appropriately shifted versions of the Δ amplitudes that conserve vector currents, so

$$\Delta_5^{\lambda\mu\nu} = \frac{1}{3}[\Delta^{\lambda\mu\nu}(a = \frac{-1}{2}(k_1 - k_2), k_1, k_2) + \Delta^{\mu\nu\lambda}(a = \frac{-1}{2}(k_2 + q), k_2, -q)$$
$$+ \Delta^{\nu\lambda\mu}(a = \frac{1}{2}(q + k_1), -q, k_1)] \ .$$

We see that

$$q_\lambda \Delta_5^{\lambda\mu\nu} = \frac{1}{3}q_\lambda \Delta^{\lambda\mu\nu}(a, k_1, k_2) = \frac{i}{6\pi^2}\varepsilon^{\mu\nu\lambda\sigma}k_{1\lambda}k_{2\sigma} \ .$$

The anomaly is spread evenly over the three vertices.

7. Define the fermionic measure $D\psi$ in (16) carefully by going to Euclidean space. Calculate the Jacobian upon a chiral transformation and derive the anomaly. [Hint: For help, see K. Fujikawa, *Phys. Rev. Lett.* 42: 1195, 1979.]

Solution:

We will proceed somewhat differently from Fujikawa's original paper. First, we will use the two-component spinor notation instead of the 4-component Dirac notation. Second, we will stay in Minkowski signature for most of the discussion and rotate to Euclidean space only to compute an integral. For this problem we use the metric convention $\eta = (-, +, +, +)$.

Consider the Lagrangian for one massless 2-component spinor ψ charged under a $U(1)$ gauge symmetry:

$$\mathcal{L} = i\psi_{\dot{a}}^\dagger \bar{\sigma}^{\mu\dot{a}a}(\partial_\mu - igA_\mu)\psi_a - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

The covariant derivative $D_\mu\psi \equiv (\partial_\mu - igA_\mu)\psi$ is present to preserve gauge invariance. For later convenience, we define the following notation:

$$v_{a\dot{a}} \equiv \sigma_{a\dot{a}}^\mu v_\mu \ , \ \bar{v}^{\dot{a}a} \equiv \bar{\sigma}^{\mu\dot{a}a}v_\mu \quad \text{for any Lorentz vector } v^\mu$$

Consider the $U(1)$ gauge transformation on ψ:

$$\psi_a'(x) = e^{i\theta(x)}\psi_a(x) = \int d^4y \, e^{i\theta(x)}\delta^4(x, y)\delta_a{}^b\psi_b(y)$$

We thus read off the Jacobian that defines this transformation:

$$J_a{}^b(x, y) \equiv \frac{\delta \psi_a'(x)}{\delta \psi_b(y)} = e^{i\theta(x)} \delta^4(x, y) \delta_a{}^b$$

For commuting numbers, the coordinate transformation $x' = Jx$ implies $\int dx' = \int dx \det J$. For anticommuting numbers, $x' = Jx$ implies $\int dx' = \int dx \, (\det J)^{-1}$. So the above change of variables gives

$$\int \mathcal{D}\psi' = \int \mathcal{D}\psi \, (\det J)^{-1} = \int \mathcal{D}\psi \, e^{-\text{tr}\ln J}$$

Imagine putting the field theory on a lattice, so that you are comfortable with $\delta^4(x, y)$ being simply the identity matrix in spacetime. The log is then:

$$\ln J = \ln\left(e^{i\theta} \mathbb{I}\right) = \ln\left[(1 + i\theta + ...)\mathbb{I}\right] = i\theta \, \mathbb{I} + O(\theta^2)$$

Putting this into the trace, we run into the immediate problem that $\text{tr}\,\mathbb{I} = \int d^4x \, \delta^4(x, x) = \infty$, but this is really no surprise at all. When we evaluate the path integral for a free scalar field theory, for example, we get a determinant:

$$\int \mathcal{D}\phi \mathcal{D}\phi^* \, \exp\left[i \int d^4x \, \phi^* \left(\partial^2 - m^2\right) \phi\right] \propto \left[\det\left(\partial^2 - m^2\right)\right]^{-1}$$

The determinant is a product of eigenvalues, and the eigenvalues can be arbitrarily large, which causes the determinant to diverge. When we can treat the determinant as an overall factor for the path integral, we don't have to worry about this. When we need to extract physical content from the determinant, we need to regulate it. Phrased in this way, the natural way to regulate the determinant is to impose an upper cutoff on the possible size of the eigenvalues.

We are now tempted to apply the above reasoning for the spinor field: when we evaluate the path integral over the fermion ψ, we get a determinant:

$$\int \mathcal{D}\psi \mathcal{D}\psi^\dagger \, \exp\left[i \int d^4x \, i\psi^\dagger \bar{D}\psi\right] \propto \det \bar{D}$$

However, continuing this line of thought fails, because the operator $\bar{D}$ does not have eigenvalues. Recall that an eigenvalue equation for the operator $\mathcal{O}$ is $\mathcal{O}\vec{v} = \lambda\vec{v}$, where λ is just a number. That is, there exists a vector $\vec{v}$ for which applying the matrix $\mathcal{O}$ returns something proportional to that same vector $\vec{v}$.

Look at the operator $\bar{D}$, defined as $\bar{D}^{\dot{a}a} \equiv \bar{\sigma}^{\mu\dot{a}a} D_\mu$. The fact that one index is dotted while the other isn't tells you everything you need to know: the operator $\bar{D}$ acts on a vector in $SU(2)_L$ but returns a vector in $SU(2)_R$ – it can't possibly have an eigenvalue equation.

We remedy this problem by constructing operators that can have eigenvalue equations. Namely, consider the operators

$$(D\bar{D})_a{}^c \equiv D_{a\dot{a}} \bar{D}^{\dot{a}c} \text{ and } (\bar{D}D)^{\dot{a}}{}_{\dot{c}} \equiv \bar{D}^{\dot{a}a} D_{a\dot{c}}$$

143

As you can see from the indices, the operator $D\bar{D}$ acts on a vector in $SU(2)_L$ and returns a vector in $SU(2)_L$. Similarly, the operator $\bar{D}D$ acts on a vector in $SU(2)_R$ and returns a vector in $SU(2)_R$. We are thus free to regulate these operators by imposing upper cutoffs on their eigenvalue spectra.

Back to the matter at hand: we wish to compute $\operatorname{tr}\ln J = \operatorname{tr} i\theta\,\mathbb{I} = \int d^4x\, i\theta(x)\operatorname{tr}\delta^4(x,x)I_2$, where $J_a{}^b(x,y) = e^{i\theta(x)}\delta^4(x,y)\delta_a{}^b$ is the Jacobian for the transformation $\psi'_a(x) = e^{i\theta(x)}\psi_a(x)$. We wish to regulate this by cutting off the eigenvalues of the operator $D\bar{D}$, which is an operator that acts purely within $SU(2)_L$. Let $|\lambda\rangle$ be an eigenstate of $D\bar{D}$ with eigenvalue λ, meaning $(D\bar{D})|\lambda\rangle = \lambda|\lambda\rangle$. Let us now regulate:

$$\delta^4(x,x) = \lim_{y\to x}\delta^4(x,y) \equiv \lim_{y\to x}\langle x|y\rangle = \lim_{y\to x}\langle x|\left(\sum_\lambda |\lambda\rangle\langle\lambda|\right)|y\rangle = \lim_{y\to x}\sum_\lambda\langle x|\lambda\rangle\langle\lambda|y\rangle$$

$$\xrightarrow{\text{regulate}}\ \lim_{y\to x}\sum_\lambda\langle x|\lambda\rangle\langle\lambda|y\rangle\, e^{-\lambda/\Lambda^2} = \lim_{y\to x}\sum_\lambda\langle x|\, e^{-D\bar{D}/\Lambda^2}|\lambda\rangle\langle\lambda|y\rangle$$

$$= \lim_{y\to x}\langle x|e^{-D\bar{D}/\Lambda^2}\left(\sum_\lambda |\lambda\rangle\langle\lambda|\right)|y\rangle = \lim_{y\to x}\langle x|e^{-D\bar{D}/\Lambda^2}|y\rangle$$

$$= \lim_{y\to x}\langle x|e^{-D\bar{D}/\Lambda^2}\left(\frac{d^4k}{(2\pi)^4}|k\rangle\langle k|\right)|y\rangle = \int\frac{d^4k}{(2\pi)^4}\lim_{y\to x}\langle x|e^{-D\bar{D}/\Lambda^2}|k\rangle\langle k|y\rangle$$

The momentum states $|k\rangle$ in the position basis are $\langle y|k\rangle = e^{+iky}$. For any function f, we know how to write D acting on it in the position basis:

$$\langle x|D_\mu|f\rangle = \left(\frac{\partial}{\partial x^\mu} - igA_\mu(x)\right)f(x)$$

We therefore know how to write any function of D acting on any function in the position basis, so in particular:

$$\langle x|e^{-D\bar{D}/\Lambda^2}|k\rangle = e^{-(D\bar{D}/\Lambda^2)(x)}e^{ikx}$$

The argument (x) for the operator $D\bar{D}$ is there to remind you that these are derivatives with respect only to x, not y, so we can move the $\langle k|y\rangle = \langle y|k\rangle^* = e^{-iky}$ to the left of the $\langle x|e^{-D\bar{D}/\Lambda^2}|k\rangle$. Taking the limit $y\to x$, we arrive at the regulated delta function:

$$\delta^4(x,x)_{\text{reg}} = \int\frac{d^4k}{(2\pi)^4}\, e^{-ikx}e^{-D\bar{D}/\Lambda^2}e^{+ikx}$$

Now let's put the operator $D\bar{D}$ into a nicer form. Introduce the usual notation $M_{(\mu\nu)}\equiv\frac{1}{2}(M_{\mu\nu}+M_{\nu\mu})$ and $M_{[\mu\nu]}\equiv\frac{1}{2}(M_{\mu\nu}-M_{\nu\mu})$, so that any matrix can be written as the sum of its symmetric and antisymmetric parts: $M_{\mu\nu} = M_{(\mu\nu)}+M_{[\mu\nu]}$. For any function $f(x)$, we have:

$$(D\bar{D})_a{}^b f(x) = \sigma^\mu_{a\dot{a}}\bar{\sigma}^{\nu\dot{a}b}D_\mu D_\nu f(x) = (\sigma^{(\mu}_{a\dot{a}}\bar{\sigma}^{\nu)\dot{a}b} + \sigma^{[\mu}_{a\dot{a}}\bar{\sigma}^{\nu]\dot{a}b})D_\mu D_\nu f(x)$$

By Lorentz invariance, we know $\sigma^{(\mu}_{a\dot{a}}\bar{\sigma}^{\nu)\dot{a}b} = C\delta_a{}^b\eta^{\mu\nu}$, where C is some constant. For $\mu = \nu = 0$, the left-hand side is just the identity matrix $\delta_a{}^b$, while the right-hand side equals $C\delta_a{}^b\eta^{00}$.

So the constant C simply cancels the 00 piece of the metric: $\eta^{00} = -1 \implies C = -1$. To summarize:

$$\sigma^{(\mu}_{a\dot{a}}\bar{\sigma}^{\nu)\dot{a}b} = -\delta_a{}^b \eta^{\mu\nu}$$

The antisymmetric piece has no neat simplification as-is, since it is just proportional to the generator of rotations. Recalling that the commutator of covariant derivatives defines the field strength, $[D_\mu, D_\nu] = -ig F_{\mu\nu}$, we proceed:

$$
\begin{aligned}
(D\bar{D})_a{}^b f(x) &= -\delta_a{}^b \eta^{\mu\nu} D_\mu D_\nu f(x) + \sigma^{[\mu}_{a\dot{a}}\bar{\sigma}^{\nu]\dot{a}b} D_\mu D_\nu f(x) \\
&= -\delta_a{}^b D^\mu D_\mu f(x) + \frac{1}{2}\sigma^{[\mu}_{a\dot{a}}\bar{\sigma}^{\nu]\dot{a}b} [D_\mu, D_\nu] f(x) \\
&= \left(-\delta_a{}^b D^2 - \frac{1}{2} ig \, \sigma^{[\mu}_{a\dot{a}}\bar{\sigma}^{\nu]\dot{a}b} F_{\mu\nu} \right) f(x)
\end{aligned}
$$

Since this expression appears exponentiated, you may expect that we'll need to raise it to various powers. Actually we will only need the square of the term with the $F_{\mu\nu}$. For later convenience, let's simplify that term now. By Lorentz invariance, we know:

$$\left(\sigma^{[\mu}\bar{\sigma}^{\nu]}\sigma^{[\rho}\bar{\sigma}^{\sigma]} \right)_a{}^b = \delta_a{}^b \left(A \, \eta^{\rho[\mu}\eta^{\nu]\sigma} + B \, \varepsilon^{\mu\nu\rho\sigma} \right)$$

Try $\mu = \rho = 0$ and $\nu = \sigma = 1$. $\sigma^{[0}\bar{\sigma}^{1]} = \frac{1}{2}(\sigma^0\bar{\sigma}^1 - \sigma^1\bar{\sigma}^0)$. Numerically, in the Weyl basis, $\sigma^0 = \bar{\sigma}^0 = I$, and $\bar{\sigma}^1 = -\sigma^1$, so we have $\sigma^{[0}\bar{\sigma}^{1]} = \frac{1}{2}(-\sigma^1 - \sigma^1) = -\sigma^1$. Therefore, the left-hand side of the equation is $(-\sigma^1)^2 = I$. Since $\varepsilon^{0101} = 0$, the right-hand side of the equation is (suppressing the 2×2 identity matrix $\delta_a{}^b$):

$$A \, \eta^{0[0}\eta^{1]1} = \frac{1}{2}A \left(\eta^{00}\eta^{11} - \eta^{01}\eta^{10} \right) = -\frac{1}{2}A$$

So comparing the two sides of the equation, we deduce $A = -2$. Now we need to solve for B. For that, consider the case $\mu = 0, \nu = 1, \rho = 2, \sigma = 3$. The right-hand side is just $B \, \varepsilon^{0123} = B$, and the left-hand side is:

$$
\begin{aligned}
\sigma^{[0}\bar{\sigma}^{1]}\sigma^{[2}\bar{\sigma}^{3]} &= \frac{1}{2^2}(\sigma^0\bar{\sigma}^1 - \sigma^1\bar{\sigma}^0)(\sigma^2\bar{\sigma}^3 - \sigma^3\bar{\sigma}^2) = \frac{1}{4}(-\sigma^1 - \sigma^1)(-\sigma^2\sigma^3 + \sigma^3\sigma^2) \\
&= +\frac{1}{2}\sigma^1[\sigma^2, \sigma^3] = \frac{1}{2}\sigma^1(+2i\sigma^1) = iI
\end{aligned}
$$

We therefore deduce that $B = i$. We have therefore arrived at the expression

$$\left(\sigma^{[\mu}\bar{\sigma}^{\nu]}\sigma^{[\rho}\bar{\sigma}^{\sigma]} \right)_a{}^b = \delta_a{}^b \left(-2 \, \eta^{\rho[\mu}\eta^{\nu]\sigma} + i \, \varepsilon^{\mu\nu\rho\sigma} \right)$$

We remind you that what we want to calculate is

$$\delta^4(x, x)_{\text{reg}} = \int \frac{d^4 k}{(2\pi)^4} e^{-ikx} e^{-D\bar{D}/\Lambda^2} e^{+ikx}$$

where the operator $D\bar{D}$ can be written as

$$(D\bar{D})_a{}^b = -\delta_a{}^b D^2 - \frac{1}{2} ig \, \sigma^{[\mu}_{a\dot{a}}\bar{\sigma}^{\nu]\dot{a}b} F_{\mu\nu}$$

145

The property $f(\partial)e^{ikx} = e^{ikx}f(\partial+ik)$ implies $D^2e^{ikx} = e^{ikx}(D+ik)^2 = e^{ikx}(D^2+2ikD-k^2)$. Any series in powers of D^2 will inherit the same property, so we get:

$$\delta^4(x,x)_{\text{reg}} = \int \frac{d^4k}{(2\pi)^4}\, e^{-ikx}e^{-D\bar{D}/\Lambda^2}e^{+ikx}$$

$$= \int \frac{d^4k}{(2\pi)^4}\, e^{-ikx}e^{+\frac{1}{\Lambda^2}(I\,D^2+\frac{1}{2}ig\sigma^{[\mu}\bar{\sigma}^{\nu]}F_{\mu\nu})}e^{+ikx}$$

$$= \int \frac{d^4k}{(2\pi)^4}\, e^{-ikx}\,e^{+ikx}e^{+\frac{1}{\Lambda^2}(I\,(D+ik)^2+\frac{1}{2}ig\sigma^{[\mu}\bar{\sigma}^{\nu]}F_{\mu\nu})}$$

$$= \int \frac{d^4k}{(2\pi)^4}\, e^{-k^2/\Lambda^2}\,e^{\frac{1}{\Lambda^2}(D^2+2ik\cdot D+\frac{1}{2}ig\sigma^{[\mu}\bar{\sigma}^{\nu]}F_{\mu\nu})}$$

$$= \Lambda^4 \int \frac{d^4k}{(2\pi)^4}\, e^{-k^2}\,e^{\frac{1}{\Lambda^2}(D^2+2i\Lambda k\cdot D+\frac{1}{2}ig\sigma^{[\mu}\bar{\sigma}^{\nu]}F_{\mu\nu})}$$

In the last line I have simply rescaled $k \to \Lambda k$ to facilitate the limit $\Lambda \to \infty$. When expanding the exponential, we see that terms of order $1/\Lambda^{>4}$ will go to zero, so we know we only need to keep finitely many terms of the series.

Before worrying about the terms that don't go to zero (and in fact look divergent), let's remind ourselves that the original Lagrangian was $\mathcal{L} = i\psi^\dagger\bar{D}\psi$, so in the path integral we need to integrate over $\psi^\dagger$ in addition to ψ. All of the previous work was to regulate the Jacobian J from the transformation

$$\psi'_a(x) = e^{i\theta(x)}\psi_a(x) \implies J_a{}^c(x,y) = e^{i\theta(x)}\delta^4(x,y)\delta_a{}^c$$

But we also need to include the Jacobian $\tilde{J}$ from the conjugate transformation

$$\psi^{\dagger\prime\dot{a}}(x) = e^{-i\theta(x)}\psi^{\dagger\dot{a}}(x) \implies \tilde{J}^{\dot{a}}{}_{\dot{c}}(x,y) = e^{-i\theta(x)}\delta^4(x,y)\delta^{\dot{a}}{}_{\dot{c}}$$

This leads to a total change in the path integral measure:

$$\int \mathcal{D}\psi'\mathcal{D}(\psi')^\dagger = \int \mathcal{D}\psi\mathcal{D}\psi^\dagger(\det J)^{-1}(\det\tilde{J})^{-1} = \int \mathcal{D}\psi\mathcal{D}\psi^\dagger\, e^{-\text{tr}\ln J-\text{tr}\ln\tilde{J}}$$

There are two things to notice here. First, most of the manipulations carry through for $\tilde{J}$ exactly as for J, so the opposite sign for the phase kills *most* of the terms upon the addition $\text{tr}\ln J+\text{tr}\ln\tilde{J}$. The second thing to notice is that the opposite sign does not kill *all* of the terms when adding the two Jacobians.

How can this be? To regularize J, we needed the operator $(D\bar{D})$, which acts purely within $SU(2)_L$. To regularize $\tilde{J}$, we need the operator $(\bar{D}D)$, which acts purely within $SU(2)_R$. Everything carries through in exactly the same way, except that squaring $D\bar{D}$ led to the expression

$$(\sigma^{[\mu}\bar{\sigma}^{\nu]}\sigma^{[\rho}\bar{\sigma}^{\sigma]})_a{}^c = \delta_a{}^c(-2\eta^{\rho[\mu}\eta^{\nu]\sigma} + i\,\varepsilon^{\mu\nu\rho\sigma})$$

whereas squaring $\bar{D}D$ leads to the expression

$$(\bar{\sigma}^{[\mu}\sigma^{\nu]}\bar{\sigma}^{[\rho}\sigma^{\sigma]})^{\dot{a}}{}_{\dot{c}} = \delta^{\dot{a}}{}_{\dot{c}}(-2\eta^{\rho[\mu}\eta^{\nu]\sigma} - i\,\varepsilon^{\mu\nu\rho\sigma})$$

The minus sign is critical: tracing and subtracting the two leaves a nonzero piece $+4i\,\varepsilon^{\mu\nu\rho\sigma}$.

Therefore, subtracting the two regularized delta functions (recall there is a trace over each set of spinor indices sitting out front) gives:

$$\text{tr }\delta^4(x,x)_{\text{reg},J} - \text{tr }\delta^4(x,x)_{\text{reg},\tilde{J}} \;=\;$$
$$\Lambda^4\int\frac{d^4k}{(2\pi)^4}\;e^{-k^2}\left[\text{tr }e^{\frac{1}{\Lambda^2}(D^2+2i\Lambda k\cdot D+\frac{1}{2}ig\sigma^{[\mu}\bar\sigma^{\nu]}F_{\mu\nu})} - \text{tr }e^{\frac{1}{\Lambda^2}(D^2+2i\Lambda k\cdot D+\frac{1}{2}ig\bar\sigma^{[\mu}\sigma^{\nu]}F_{\mu\nu})}\right]$$

The first trace is over $SU(2)_L$ indices, and the second trace is over $SU(2)_R$ indices, which of course is the way it must be for the expression to make any sense.

Now you see what is happening here. The lowest order term (just the identity matrix) drops out. The first order term also drops out, except for the subtraction $\sigma^{[\mu}\bar\sigma^{\nu]} - \bar\sigma^{[\mu}\sigma^{\nu]}$. Since $\sigma^{[\mu}\bar\sigma^{\nu]} = \frac{1}{2}(\sigma^\mu\bar\sigma^\nu - \sigma^\nu\bar\sigma^\mu)$, the cyclic property of the trace over $SU(2)_L$ indices causes that term to be zero. Similarly, the trace over $SU(2)_R$ indices yields $\text{tr}(\bar\sigma^{[\mu}\sigma^{\nu]}) = 0$.

So expanding the exponentials yields terms that will cancel each other out, trace to zero, or will go to 0 for $\Lambda\to\infty$. As discussed, the only exception is:

$$\text{tr }(\sigma^{[\mu}\bar\sigma^{\nu]}\sigma^{[\rho}\bar\sigma^{\sigma]}) - \text{tr }(\bar\sigma^{[\mu}\sigma^{\nu]}\bar\sigma^{[\rho}\sigma^{\sigma]}) = \delta_a{}^a\,i\varepsilon^{\mu\nu\rho\sigma} + \delta^{\dot a}{}_{\dot a}\,i\varepsilon^{\mu\nu\rho\sigma} = 4i\,\varepsilon^{\mu\nu\rho\sigma}$$

So tracing over the delta functions and subtracting them, and then taking the cutoff to infinity, gives

$$\text{tr }\delta^4(x,x)_{\text{reg},J} - \text{tr }\delta^4(x,x)_{\text{reg},\tilde{J}} = \int\frac{d^4k}{(2\pi)^4}\;e^{-k^2}\frac{1}{2}\left(\frac{ig}{2}\right)^2 4i\,\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$$
$$= -\frac{1}{2}ig^2\,\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\int\frac{d^4k}{(2\pi)^4}\;e^{-k^2}$$

Now Wick rotate to evaluate the integral: $k^0 \equiv -ik^0_E \implies k^2 \equiv k_\mu k^\mu = -(k^0)^2 + \vec{k}^2 = +(k^0_E)^2 + \vec{k}^2 \equiv +k^2_E$ and $\int d^4k = -i\int d^4k_E$. We obtain for the integral

$$\int\frac{d^4k}{(2\pi)^4}\;e^{-k^2} = -i\int\frac{d^4k_E}{(2\pi)^4}\;e^{-k^2_E} = -i\left(\int_{-\infty}^{\infty}\frac{dk}{2\pi}\;e^{-k^2}\right)^4 = -i\left(\frac{1}{2\sqrt{\pi}}\right)^4 = \frac{-i}{16\pi^2}$$

So our final expression for the delta functions is

$$\text{tr }\delta^4(x,x)_{\text{reg},J} - \text{tr }\delta^4(x,x)_{\text{reg},\tilde{J}} \;=\; -\frac{g^2}{32\pi^2}\,\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$$

Now recall $\text{tr }\ln J = \int d^4x\,i\theta(x)\,\text{tr }\delta^4(x,x)_{\text{reg}}$, so

$$\text{tr }\ln J + \text{tr }\ln\tilde{J} = i\int d^4x\,\theta(x)\left(-\frac{g^2}{32\pi^2}\,\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\right)$$

We have finally managed to compute the change in the path integral measure. Under the change of integration variables

$$\psi'_a(x) = e^{i\theta(x)}\psi_a(x) \; , \; \psi^{\dagger'\dot{a}}(x) = e^{-i\theta(x)}\psi^{\dagger\dot{a}}(x)$$

the path integral measure changes as

$$\int \mathcal{D}\psi'\mathcal{D}(\psi')^{\dagger} = \int \mathcal{D}\psi\mathcal{D}\psi^{\dagger}\, e^{i\int d^4x\,\theta(x)\left(+\frac{1}{32\pi^2}g^2\,\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\right)}$$

We see that the measure is not invariant under this change of variables (in other words, the Jacobian is not 1).

Recall the original Lagrangian:

$$\mathcal{L} = i(\psi')^{\dagger}_{\dot{a}}\bar{\sigma}^{\mu\dot{a}a}(\partial_\mu - igA_\mu)\psi'_a - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

where now we have primed the fields to follow the order in which we have defined the change of variables all along. Under the infinitesimal version of this change of variables, the Lagrangian becomes $\mathcal{L}' = \mathcal{L} + \delta\mathcal{L}$, where

$$\delta\mathcal{L} = i\psi^{\dagger}_{\dot{a}}\bar{\sigma}^{\mu\dot{a}a}(i\,\partial_\mu\theta)\psi_a = +\theta(x)\partial_\mu\left(\psi^{\dagger}_{\dot{a}}\bar{\sigma}^{\mu\dot{a}a}\psi_a\right) + (\text{total derivative})$$

Finally, we are ready to state the result. The infinitesimal transformation $\psi \to \psi + \delta\psi$, $\psi^{\dagger} \to \psi^{\dagger} + \delta\psi^{\dagger}$ with $\delta\psi_a(x) = i\theta(x)\,\psi_a(x)$, $\delta\psi^{\dagger\dot{a}} = -i\theta(x)\,\psi^{\dagger\dot{a}}$ changes the path integral as:

$$\int \mathcal{D}\psi\mathcal{D}\psi^{\dagger}\, e^{i\int d^4x\,\mathcal{L}} \to \int \mathcal{D}\psi\mathcal{D}\psi^{\dagger}\, e^{i\int d^4x\,\mathcal{L} + i\int d^4x\,\theta(x)\left[\partial_\mu\left(\psi^{\dagger}_{\dot{a}}\bar{\sigma}^{\mu\dot{a}a}\psi_a\right) + \frac{1}{32\pi^2}g^2\,\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}\right]}$$

Just as the integral $\int_{-\infty}^{\infty} dx\, f(x)$ is not a function of x, the path integral $\int \mathcal{D}\psi\mathcal{D}\psi^{\dagger}\, e^{iS(\psi,\psi^{\dagger})}$ is not a function of ψ or $\psi^{\dagger}$. Therefore the path integral should remain completely unchanged under the infinitesimal transformation described above. Moreover, since this transformation is for an arbitrary infinitesimal function of spacetime $\theta(x)$, we conclude that

$$\partial_\mu\left(\psi^{\dagger}_{\dot{a}}\bar{\sigma}^{\mu\dot{a}a}\psi_a\right) = -\frac{1}{32\pi^2}g^2\,\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$$

This entire analysis was for one complex Weyl spinor. Suppose we introduce a second complex Weyl spinor ξ. The Lagrangian density is

$$\mathcal{L} = i\psi^{\dagger}_{\dot{a}}\bar{\sigma}^{\mu\dot{a}a}(\partial_\mu - igA_\mu)\psi_a + i\xi^{\dagger}_{\dot{a}}\bar{\sigma}^{\mu\dot{a}a}(\partial_\mu - igA_\mu)\xi_a - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

This Lagrangian exhibits a classical $U(1)$ gauge symmetry

$$\begin{pmatrix} \psi_a \\ \xi_a \end{pmatrix} \to \begin{pmatrix} e^{i\theta(x)}\psi_a \\ e^{i\theta(x)}\xi_a \end{pmatrix}$$

148

with the associated Noether current

$$j^\mu = \psi^\dagger_{\dot a}\bar\sigma^{\mu\dot a a}\psi_a + \xi^\dagger_{\dot a}\bar\sigma^{\mu\dot a a}\xi_a$$

Classically Noether's theorem says $\partial_\mu j^\mu = 0$, but we have just derived that ψ and ξ each contributes a nonzero contribution to this classical conservation law, so that at the quantum level we have

$$\partial_\mu j^\mu = -\frac{1}{16\pi^2}g^2\,\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$$

Thus this theory does not actually have the $U(1)$ gauge symmetry, even though it appears as though it did at the classical level. That is the statement of the chiral anomaly.

The Lagrangian also exhibits a different classical $U(1)$ gauge symmetry

$$\begin{pmatrix}\psi_a\\\xi_a\end{pmatrix} \to \begin{pmatrix}e^{i\theta(x)}\psi_a\\e^{-i\theta(x)}\xi_a\end{pmatrix}$$

with the associated Noether current

$$j^\mu = \psi^\dagger_{\dot a}\bar\sigma^{\mu\dot a a}\psi_a - \xi^\dagger_{\dot a}\bar\sigma^{\mu\dot a a}\xi_a$$

Note the relative minus sign between the two terms: the contributions to $\partial_\mu j^\mu$ cancel exactly. Thus the theory truly does have this $U(1)$ gauge symmetry, even quantum mechanically.

To compare this to the usual treatment in terms of Dirac spinors, let us package the 2-component fields ψ_a and ξ_a into a 4-component Dirac spinor:

$$\Psi \equiv \begin{pmatrix}\psi_a\\\xi^{\dagger\dot a}\end{pmatrix}$$

The $U(1)$ transformation

$$\begin{pmatrix}\psi_a\\\xi_a\end{pmatrix} \to \begin{pmatrix}e^{i\theta(x)}\psi_a\\e^{i\theta(x)}\xi_a\end{pmatrix} \implies \begin{pmatrix}\psi_a\\\xi^{\dagger\dot a}\end{pmatrix} \to \begin{pmatrix}e^{i\theta(x)}\psi_a\\e^{-i\theta(x)}\xi^{\dagger\dot a}\end{pmatrix}$$

is the transformation $\Psi \to e^{i\theta(x)\gamma_5}\Psi$. The left-handed and right-handed components of the Dirac spinor Ψ are transformed oppositely, which is why this $U(1)$ transformation is called chiral. On the other hand, the $U(1)$ transformation

$$\begin{pmatrix}\psi_a\\\xi_a\end{pmatrix} \to \begin{pmatrix}e^{i\theta(x)}\psi_a\\e^{-i\theta(x)}\xi_a\end{pmatrix} \implies \begin{pmatrix}\psi_a\\\xi^{\dagger\dot a}\end{pmatrix} \to \begin{pmatrix}e^{i\theta(x)}\psi_a\\e^{i\theta(x)}\xi^{\dagger\dot a}\end{pmatrix}$$

is the transformation $\Psi \to e^{i\theta(x)}\Psi$. The current associated with this transformation is the usual electromagnetic current, which is indeed conserved.

8. Compute the pentagon anomaly by Feynman diagrams in order to check remark 6 in the text. In other words, determine the coefficient c in $\partial_\mu J_5^\mu = ... + c\varepsilon^{\mu\nu\lambda\sigma}\mathrm{tr}A_\mu A_\nu A_\lambda A_\sigma$.

Solution:

The pentagon diagram has 5 fermion propagators, so the integral goes as $\int_\Lambda d^4k \frac{1}{k^5} \sim \frac{1}{\Lambda}$, which is convergent as $\Lambda \to \infty$ and thereby naively appears not to contribute to the anomaly. However, gauge invariance relates the pentagon diagram to the triangle and box diagrams, so that only the set of all three is physically meaningful. Demanding that the box diagram conserves vector currents then generates a pentagon anomaly in the form

$$\partial_\mu J_5^{a\mu} = ... + \frac{1}{4\pi^2}\varepsilon^{\mu\nu\lambda\sigma}\mathrm{tr}\{(T_5^a)[\tfrac{1}{4}v_{\mu\nu}v_{\lambda\sigma} + \tfrac{1}{12}a_{\mu\nu}a_{\lambda\sigma}]$$
$$+ i\tfrac{2}{3}(a_\mu a_\nu v_{\lambda\sigma} + v_{\mu\nu}a_\lambda a_\sigma + a_\mu v_{\nu\lambda}a_\sigma) - \tfrac{8}{3}a_\mu a_\nu a_\lambda a_\sigma]\} \ .$$

Here we have split up the gauge field into vector and axial vector components, $A = \slashed{v} + \gamma^5\slashed{a}$, and defined the corresponding field strengths $v_{\mu\nu} = \partial_\mu v_\nu - \partial_\nu v_\mu - i[v_\mu, v_\nu] - i[a_\mu, a_\nu]$ and $a_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu - i[v_\mu, a_\nu] - i[a_\mu, v_\nu]$. The matrix T_5^a is the generator associated with the axial coupling to fermions: $a_\mu = a_\mu^a T_5^a$.

See W. A. Bardeen, "Anomalous Ward Identities in Spinor Field Theories," Phys. Rev. Vol. 184 No. 5, 25 Aug 1969 for further details.

V Field Theory and Collective Phenomena

V.1 Superfluids

2. To confine the superfluid in an external potential $W(\vec{x})$ we would add the term

$$-W(\vec{x})\varphi^\dagger(\vec{x}, t)\varphi(\vec{x}, t)$$

to (1). Derive the corresponding equation of motion for φ. The equation, known as the Gross-Pitaevski equation, has been much studied in recent years in connection with the Bose-Einstein condensate.

$$\mathcal{L} = i\varphi^\dagger\partial_0\varphi - \frac{1}{2m}\partial_i\varphi^\dagger\partial_i\varphi - g^2(\varphi^\dagger\varphi - \bar{\rho})^2 \qquad (1)$$

Solution:

Adding the term $\mathcal{L}_{\mathrm{pot}} = -W(\vec{x})\varphi^\dagger\varphi$ implies the addition of a term $-W(\vec{x})\varphi$ to the equation of motion found from varying $\mathcal{L}$ with respect to $\varphi^\dagger$.

V.2 Euclid, Boltzmann, Hawking, and Field Theory at Finite Temperature

1. Study the free field theory $\mathcal{L} = \frac{1}{2}(\partial\varphi)^2 - \frac{1}{2}m^2\varphi^2$ at finite temperature and derive the Bose-Einstein distribution.

Solution:

One way to derive the Bose-Einstein distribution is to show that average quantities are computed with the correct probability distribution.[11]

We first compute the thermodynamic partition function for the relativistic free Bose gas, given by:

$$\mathcal{Z} = \int_{\phi(0,\vec{x})=\phi(\beta,\vec{x})} \mathcal{D}\phi \, e^{-\int_0^\beta d\tau \int d^3x \, \mathcal{L}_E} , \quad \mathcal{L}_E = \frac{1}{2}\partial_\mu\phi \, \partial_\mu\phi + \frac{1}{2}M^2\phi^2 .$$

The path integral is Gaussian and therefore can be computed exactly:

$$\mathcal{Z} = [\det(-\partial_E^2 + M^2)]^{-1/2} = e^{-\frac{1}{2}\mathrm{tr}\ln(-\partial_E^2 + M^2)} .$$

Here we have dropped an overall constant, which corresponds to an additive constant in the argument of the exponent. The trace is the sum of eigenvalues of the operator $-\partial_E^2 + M^2$, which is diagonalized in Matsubara frequency / momentum space. So we have[12]

$$\ln \mathcal{Z} = -\frac{1}{2} \sum_{n=-\infty}^{\infty} \int \frac{d^3k V}{(2\pi)^3} \ln(\omega_n^2 + \vec{k}^2 + M^2) , \quad \omega_n = \frac{2\pi n}{\beta} .$$

Consider the sum

$$f(x) \equiv \sum_{n=1}^{\infty} \ln(n^2 + x^2) .$$

While this sum diverges, its derivative is finite and can be computed:

$$f'(x) = \sum_{n=1}^{\infty} \frac{2x}{n^2 + x^2} = -\pi - \frac{1}{x} + \frac{2\pi}{1 - e^{-2\pi x}} .$$

Using $f(x) = \int dx \, f'(x)$, we find:

$$\sum_{n=-\infty}^{\infty} \ln(n^2 + x^2) = 2\pi x + 2\ln\left(1 - e^{-2\pi x}\right) .$$

---

[11]We follow p. 394 of "Lattice Gauge Theories - An Introduction," 3rd Ed, by Heinz J. Rothe.

[12]We have implicitly put the system into a box to quantize the spatial momenta and then taken the continuum limit $\sum_k \to \int \frac{d^3k}{(2\pi)^3} V$, where V is the volume of the box. This treatment misses the formation of the Bose-Einstein condensate.

Therefore we can compute the sum over Matsubara frequencies to obtain, up to an irrelevant additive constant,

$$\ln \mathcal{Z} = -V \int \frac{d^3k}{(2\pi)^3} \ln\left(1 - e^{-\beta\varepsilon(\vec{k})}\right) \ , \quad \varepsilon(\vec{k}) \equiv \sqrt{\vec{k}^2 + M^2} \ .$$

The average energy is given by $\langle E \rangle = -\frac{\partial}{\partial\beta} \ln \mathcal{Z}$, so that the average energy per unit volume is given by:

$$\left\langle \frac{E}{V} \right\rangle = \frac{\partial}{\partial\beta} \int \frac{d^3k}{(2\pi)^3} \ln\left(1 - e^{-\beta\varepsilon(\vec{k})}\right)$$
$$= \int \frac{d^3k}{(2\pi)^3} \varepsilon(\vec{k}) \frac{1}{e^{\beta\varepsilon(\vec{k})} - 1} \ .$$

Thus the average energy is computed with the probability distribution

$$n(\varepsilon) = \frac{1}{e^{\beta\varepsilon} - 1}$$

which is the Bose-Einstein distribution.

2. It probably does not surprise you that for fermionic fields the periodic boundary condition (6) is replaced by an antiperiodic boundary condition $\psi(\vec{x}, 0) = -\psi(\vec{x}, \beta)$ in order to reproduce the results of chapter II.5. Prove this by looking at the simplest fermionic functional integral. [Hint: The clearest exposition of this satisfying fact may be found in appendix A of R. Dashen, B. Hasslacher, and A. Neveu, Phys. Rev. D12: 2443, 1975.]

Solution:

Consider the case of free field theory in $(0 + 1)$-spacetime dimensions (that is, the quantum mechanics of a fermionic harmonic oscillator).

The Lagrangian is $L = \bar{\psi}(t)(i\partial_t - m)\psi(t)$. The system has two possible single-particle energy states, ε_0 and ε_1, corresponding to whether or not a fermion is present. The energy difference $\varepsilon_1 - \varepsilon_0 = m$ is fixed, while the energies themselves are determined only up to an overall additive constant: $\varepsilon_0 = E_0 - \frac{1}{2}m$ and $\varepsilon_1 = E_0 + \frac{1}{2}m$.

We can now compute the partition function:

$$\mathcal{Z} \equiv \mathrm{tr}(e^{-iHT}) = \sum_{n=0}^{1} e^{-i\varepsilon_n T} = 2e^{-iE_0 T} \cos\left(\frac{mT}{2}\right) \ .$$

Any path integral representation of the partition function must reproduce this basic result.

The standard tricks give us the path integral representation

$$\mathcal{Z} = \int_{\pm \text{BC}} \mathcal{D}\psi \mathcal{D}\bar{\psi} \, e^{i \int_0^T dt \, \bar{\psi}(t)(i\partial_t - m)\psi(t)}$$

where the $\pm$BC indicates periodic or antiperiodic boundary conditions, respectively: $\psi(t + T) = \pm\psi(t)$.

The integral is gaussian and may be performed explicitly:

$$\mathcal{Z} = C \det(i\partial_t - m)$$

where C is a normalization factor that is determined by the overall additive constant in the energy:

$$\text{tr}(e^{-iH(m=0)T}) = \int_{\pm \text{BC}} \mathcal{D}\psi \mathcal{D}\bar{\psi} \, e^{i \int_0^T dt \, \bar{\psi}(t) i\partial_t \psi(t)} = C \det(i\partial_t) \equiv 2e^{-iE_0 T}$$

$$\implies C = \frac{2e^{-iE_0 T}}{\det(i\partial_t)} \ .$$

Therefore, the path integral calculation gives a partition function

$$\mathcal{Z} = 2e^{-iE_0 T} \frac{\det(i\partial_t - m)}{\det(i\partial_t)} \ .$$

The determinant of an operator is the product of its eigenvalues. The eigenvalue equation $(i\partial_t - m)\varphi_n(t) = -\omega_n(m)\varphi_n(t)$ subject to $\pm$ boundary conditions implies the eigenvalues $(n = 0, \pm 1, \pm 2, ...)$:

$$\omega_n(m) = \begin{cases} \frac{2n\pi}{T} + m & \text{(periodic)} \\[2ex] \frac{(2n+1)\pi}{T} + m & \text{(antiperiodic)} \end{cases}$$

Since

$$\frac{\det(i\partial_t - m)}{\det(i\partial_t)} = \prod_{n=-\infty}^{\infty} \frac{\omega_n(m)}{\omega_n(0)}$$

we have a choice between two infinite products. Since

$$\prod_{n=-\infty}^{\infty} \left(1 + \frac{x}{2n+1}\right) = \cos\left(\frac{\pi x}{2}\right)$$

we deduce that only antiperiodic boundary conditions will give us the correct answer:

$$\frac{\det(i\partial_t - m)}{\det(i\partial_t)}\bigg|_{-\text{BC}} = \prod_{n=-\infty}^{\infty} \frac{\frac{(2n+1)\pi}{T} + m}{\frac{(2n+1)\pi}{T}} = \prod_{n=-\infty}^{\infty} \left(1 + \frac{mT}{(2n+1)\pi}\right) = \cos\left(\frac{mT}{2}\right) \ .$$

153

By the usual analytic continuation $T = -i\beta$, we conclude that fermions at finite temperature obey antiperiodic boundary conditions.

3. It is interesting to consider quantum field theory at finite density, as may occur in dense astrophysical objects or in heavy ion collisions. (In the previous chapter we studied a system of bosons at finite density and zero temperature.) In statistical mechanics we learned to go from the partition function to the grand partition function $Z = \text{tr}\, e^{-\beta(H-\mu N)}$, where a chemical potential μ is introduced for every conserved particle number N. For example, for noninteracting relativistic fermions, the Lagrangian is modified to $\mathcal{L} = \bar{\psi}(i\,\partial\!\!\!/ - m)\psi + \mu\bar{\psi}\gamma^0\psi$. Note that finite density, as well as finite temperature, breaks Lorentz invariance. Develop the subject of quantum field theory at finite density as far as you can.

Solution:

First let us clarify that finite density breaks Lorentz invariance because the term $\bar{\psi}\gamma^\mu\psi$ is a Lorentz 4-vector, and we add to the Lagrangian only the $\mu = 0$ component of that term, $\bar{\psi}\gamma^0\psi = \psi^\dagger\psi$.

To work with a theory at nonzero temperature $T \equiv \beta^{-1}$, take the Minkowski theory and compactify the imaginary time direction: $t \to i\tau$, $0 \le \tau \le \beta$. The gamma matrices of $SO(3,1)$ satisfy the relation $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$, so that $(\gamma^0)^2 = +1$ while $(\gamma^1)^2 = (\gamma^2)^2 = (\gamma^3)^2 = -1$. The gamma matrices of $SO(4)$ satisfy $\{\gamma_E^\mu, \gamma_E^\nu\} = 2\delta^{\mu\nu}$, so that $(\gamma_E^0)^2 = (\gamma_E^1)^2 = (\gamma_E^2)^2 = (\gamma_E^3)^2 = +1$, where the subscript E stands for "Euclidean." We can therefore write

$$\gamma_E^0 = \gamma^0\,, \quad \gamma_E^i = i\gamma^i$$

for the Euclidean gamma matrices[13]. The Dirac operator in Euclidean space is

$$\partial\!\!\!/ = \gamma^\mu\partial_\mu = \gamma^0\partial_t + \gamma^i\partial_i = \gamma_E^0(-i\partial_\tau) + (-i\gamma_E^i)\partial_i = -i(\gamma_E^0\partial_\tau + \gamma_E^i\partial_i) \equiv -i\partial\!\!\!/_E\,.$$

The theory we will work with has the Euclidean action $S_E = \int_\beta d^4x\, \mathcal{L}_E$, where we have defined $\int_\beta d^4x \equiv \int_0^\beta dx^0 \int d^3x$, and the Lagrangian

$$\mathcal{L}_E = \bar{\psi}(\partial\!\!\!/_E + m)\psi + \mu\bar{\psi}\gamma_E^0\psi\,.$$

The thermal two-point function $\mathcal{S}(x) \equiv \langle \psi(x)\bar{\psi}(0) \rangle$ is defined as the inverse of the operator $\partial\!\!\!/_E + m + \mu\gamma_E^0$:

$$\partial\!\!\!/_E + m + \mu\gamma_E^0\,\mathcal{S}(x) = I\delta^4(x)$$

where I is the 4×4 identity matrix in Dirac spinor space. Multiply this equation by $e^{-ip_\mu x_\mu}$ and integrate $\int_\beta d^4x$ to get

$$(+i\,p\!\!\!/_E + m + \mu\gamma_E^0)\mathcal{S}(p) = I$$

---

[13]We continue to label the spacetime indices by 0,1,2,3 in Euclidean spacetime, instead of the often-used convention 1,2,3,4.

where we have defined the Fourier transform

$$\mathcal{S}(p) \equiv \int_\beta d^4 x \, e^{-i p_\mu x_\mu} \mathcal{S}(x) \; .$$

We therefore have the thermal two-point function

$$\mathcal{S}(x) = \sum_{n=-\infty}^{\infty} \frac{1}{\beta} \int \frac{d^3 p}{(2\pi)^3} \frac{-i(\omega_n + i\mu)\gamma_E^0 - i p_i \gamma_E^i + m}{(\omega_n + i\mu)^2 + \vec{p}^2 + m^2} e^{+i(\omega_n \tau + \vec{p} \cdot \vec{x})}$$

where we have defined $p_0 = \omega_n \equiv \frac{\pi}{\beta}(2n+1)$ and summed instead of integrated because the $\tau \equiv x^0$ direction is compact. Everything proceeds as in the case for which $\mu = 0$, except with the replacement $p_0 \to p_0 + i\mu$.

V.3   Landau-Ginzburg Theory of Critical Phenomena

1. Another important critical exponent γ is defined by saying that the susceptibility $\chi \equiv (\partial M / \partial H)|_{H=0}$ diverges as $\sim 1/|T - T_c|^\gamma$ as T approaches T_c. Determine γ in Landau-Ginzburg theory. [Hint: Instructively, there are two ways of doing it: (a) Add $-\vec{H} \cdot \vec{M}$ to (1) for $\vec{M}$ and $\vec{H}$ constant in space and solve for $\vec{M}(\vec{H})$. (b) Calculate the susceptibility function $\chi_{ij}(x - y) \equiv [\partial M_i(x)/\partial H_j(y)]|_{H=0}$ and integrate over space.]

Solution:

a) Adding $-\vec{H} \cdot \vec{M}$ to (1) gives the free energy

$$G = (a\vec{M}^2 - \vec{H} \cdot \vec{M})V + O[(\vec{M}^2)^2]$$

This is minimized for $\vec{M} = \frac{1}{a}\vec{H}$, so the susceptibility is $\chi = \partial M / \partial H = \frac{1}{a} \sim 1/(T - T_c)$, which implies $\gamma = 1$ in Landau-Ginzburg theory.

b) Starting from (3) on p. 293, we have the susceptibility function

$$\chi_{ij}(\vec{x}) \equiv \left(\frac{\partial M_i(\vec{x})}{\partial H_j(\vec{y})} \right)_{H=0} = \delta_{ij} \frac{e^{-\sqrt{a}|\vec{x}-\vec{y}|}}{4\pi|\vec{x}-\vec{y}|}$$

Stripping off the δ_{ij} and integrating over space gives the susceptibility

$$\chi = \int d^3 x \, \frac{e^{-\sqrt{a}|\vec{x}|}}{4\pi|\vec{x}|} = \int_0^\infty dr \, r \, e^{-\sqrt{a}\, r} = \frac{1}{a} \sim \frac{1}{T - T_c}$$

Again we find $\gamma = 1$.

V.4 Superconductivity

1. Vary (1) to obtain the equation for A and determine the London penetration length more carefully.

$$(1) \qquad \mathcal{F} = \frac{1}{4}F_{ij}^2 + |D_i\varphi|^2 + a|\varphi|^2 + \frac{b}{2}|\varphi|^4 + \dots$$

Solution:

Expanding out the field strength $F_{ij} = \partial_i A_j - \partial_j A_i$ gives

$$\frac{1}{4}F_{ij}F_{ij} = \frac{1}{2}A_i\left(-\delta_{ij}\vec{\partial}^2 + \partial_i\partial_j\right)A_j \qquad (i)$$

and the covariant derivative $D_i\varphi = \partial_i\varphi - 2ieA_i\varphi$ implies

$$(D_i\varphi)^\dagger D_i\varphi = \partial_i\varphi^\dagger\partial_i\varphi + 2ieA_i(\varphi^\dagger\partial_i\varphi - \varphi\partial_i\varphi^\dagger) + (2e)^2\varphi^\dagger\varphi A_i A_i \ . \qquad (ii)$$

Varying equation (i) with respect to A_i gives

$$\delta\left(\frac{1}{4}F_{ij}F_{ij}\right) = \delta A_i\left(-\delta_{ij}\vec{\partial}^2 + \partial_i\partial_j\right)A_j$$

and varying equation (ii) with respect to A_i gives

$$\delta\left[(D_i\varphi)^\dagger D_i\varphi\right] = \delta A_i\left[2ie(\varphi^\dagger\partial_i\varphi - \varphi\partial_i\varphi^\dagger) + 2(2e)^2\varphi^\dagger\varphi A_i\right] \ .$$

The equation of motion found from setting $\delta\mathcal{F} = \mathcal{F}[A + \delta A] - \mathcal{F}[A] = 0$ is therefore

$$\left(-\delta_{ij}\vec{\partial}^2 + \partial_i\partial_j + \frac{m_\gamma^2}{\langle\varphi^\dagger\varphi\rangle}\varphi^\dagger\varphi\,\delta_{ij}\right)A_j = -2ie(\varphi^\dagger\partial_i\varphi - \varphi\partial_i\varphi^\dagger)$$

where we have defined the mass of the photon inside the superconductor via $m_\gamma^2 \equiv 2(2e)^2\langle\varphi^\dagger\varphi\rangle$. The extra factor of two is there because the mass term for a real vector boson is $\mathcal{F} = \frac{1}{2}m_A^2 A_i A_i$.

Choose the gauge $\partial_j A_j = 0$. When φ assumes its vacuum expectation value, we have $(-\vec{\partial}^2 + m_\gamma^2)A_i = 0$. The physically relevant solution is the decaying exponential. Substitute the form $A_i(\vec{x}) = C\,e^{-\vec{k}\cdot\vec{x}}$ to get

$$(-\vec{k}^2 + m_\gamma^2)C = 0 \implies |\vec{k}| = m_\gamma \ .$$

We find an exponential decay $A_i(r) \sim e^{-m_\gamma r}$. The London penetration depth is defined as the length scale over which the field penetrates into the material, meaning

$$\ell_L = \frac{1}{m_\gamma} = \frac{1}{\sqrt{2(2e)^2 v^2}} = \frac{1}{2\sqrt{2}\,ev}$$

where $v \equiv \sqrt{\langle\varphi^\dagger\varphi\rangle}$.

2. Determine the coherence length more carefully.

Solution:

Use the free energy from problem V.4.1, but this time vary $\varphi^\dagger$ to get an equation of motion for φ. The relevant part of the free energy is

$$\mathcal{F} = \partial_i \varphi^\dagger \partial_i \varphi + 2ie A_i (\varphi^\dagger \partial_i \varphi - \varphi \partial_i \varphi^\dagger) + \frac{m_\gamma^2}{2v^2} \varphi^\dagger \varphi A_i A_i + a \varphi^\dagger \varphi + \frac{b}{2}(\varphi^\dagger \varphi)^2 \ .$$

If we choose the gauge $\partial_i A_i = 0$, then we can integrate by parts to move the derivative off of $\varphi^\dagger$ and onto φ. The variation $\delta \mathcal{F} = \mathcal{F}[\varphi^\dagger + \delta\varphi^\dagger] - F[\varphi^\dagger]$ immediately gives the equation of motion for φ:

$$\left(-\vec{\partial}^2 + 4ie\, \vec{A} \cdot \vec{\partial} + \frac{m_\gamma^2}{2v^2} \vec{A}^2 + a + b\, \varphi^\dagger \varphi \right) \varphi = 0 \ .$$

Deep inside the superconductor (or more accurately, at depths much greater than ℓ_L derived in the previous problem), we can set the vector potential to zero. Assuming that the self-interaction term governed by the parameter b is negligible, we obtain

$$(-\vec{\partial}^2 + a)\varphi = 0$$

Recall that for this entire discussion to work, we must have $a < 0$ (the analog of the "negative mass squared" instability from chapter IV.) We therefore determine an oscillating solution $\varphi(r) \sim e^{\pm ir/\ell_\varphi}$, where

$$\ell_\varphi = \frac{1}{\sqrt{-a}}$$

is the coherence length, defined as the characteristic length scale over which φ varies.

V.6 Solitons

3. Compute the mass of the kink by the brute force method and check the result from the Bogomol'nyi inequality.

Solution:

The mass of the kink is given on p. 305 as

$$M = \frac{\mu^2}{\lambda}\mu \int dy \left[\frac{1}{2}\left(\frac{df}{dy}\right)^2 + \frac{1}{4}\left(f^2 - 1\right)^2 \right]$$

where $f(y) \equiv \varphi(x)/v$ and $y \equiv \mu x$. Setting $\delta M \equiv M[f + \delta f] - M[f] = 0$ gives the equation $f''(y) = [f(y)^2 - 1]f(y)$. Mathematica can solve this equation using the command

$$\mathrm{DSolve}[\{\mathrm{f\ ''[y]} == (\mathrm{f[y]^2 - 1})\mathrm{f[y]}, \mathrm{f\ '[\infty]} == 0\}, \mathrm{f[y]}, \mathrm{y}]$$

which yields the solution

$$f(y) = \tanh\left(\frac{y}{\sqrt{2}}\right) .$$

Putting this solution into M and integrating from $-\infty < y < \infty$, we get the result

$$M = \frac{2\sqrt{2}}{3}\frac{\mu^2}{\lambda}\mu .$$

The Bogomol'nyi inequality is

$$M \geq \frac{2\sqrt{2}}{3}\frac{\mu^2}{\lambda}\mu|Q|$$

so the kink is at precisely the lower bound of the inequality. This implies $|Q| \leq 1$, which is another way of showing that the kink has topological charge $Q = 1$ (using the convention for which the antikink has charge $Q = -1$).

V.7 Vortices, Monopoles, and Instantons

1. Explain the relation between the mathematical statement $\Pi_0(S^0) = Z_2$ and the physical result that there are no kinks with $|Q| \geq 2$.

Solution:

The quantity $\Pi_0(S^0)$ counts the number of topologically inequivalent mappings of spatial S^0 into the group manifold $M = S^0$. Since $\Pi_0(S^0) = \mathbb{Z}_2$, and since the group $\mathbb{Z}_2$ has only two elements, there are only two topologically inequivalent kink configurations. These are the kink and antikink, with $Q = +1$ and $Q = -1$ respectively. Therefore, there is no solution with a nonzero Q for which $|Q| \neq 1$.

2. In the vortex, study the length scales characterizing the variation of the fields φ and A. Estimate the mass of the vortex.

Solution:

The mass of a time-independent configuration of the theory is

$$M = \int d^2x \left[(D_i\varphi)^\dagger D_i\varphi + \lambda(\varphi^\dagger\varphi - v^2)^2 + \frac{1}{4}F_{ij}F_{ij} \right]$$

where $D_i\varphi = \partial_i\varphi - ieA_i\varphi$. For future reference, let us collect the mass dimensions of all quantities in this expression. In (2+1) spacetime dimensions, we have: $[\varphi] = [v] = [A] = [e] = +\frac{1}{2}$ and $[\lambda] = +1$. To check this, note that $[\partial] = +1$ must have the same dimension as $[eA] = [e] + [A] = +\frac{1}{2} + \frac{1}{2} = +1$ ✓. Also, $[F] = [\partial A] = +1 + \frac{1}{2} = +\frac{3}{2}$, so that $[F^2] = +3$, which is indeed the mass dimension of the Lagrangian density in (2+1) dimensions. Finally, $[\lambda\varphi^4] = [\lambda] + 4[\varphi] = +1 + 2 = +3$, which is again the correct mass dimension of the Lagrangian.

Propose the ansatz

$$\varphi(r,\theta) = ve^{i\theta}f(r) , \quad A_i(r,\theta) = \frac{1}{e}\partial_i\theta\, g(r)$$

with $f(r)$ and $g(r)$ dimensionless functions that equal 1 at $r = \infty$. Using the relations

$$\partial_i r = \frac{x_i}{r} , \quad \partial_i\theta = \frac{1}{r}(\cos\theta\, \delta_{iy} - \sin\theta\, \delta_{ix})$$

we can now perform the requisite algebra to obtain each part of the Lagrangian. The derivative of the scalar is

$$\partial_i\varphi = \frac{v}{r}e^{i\theta}\left[i(\cos\theta\, \delta_{iy} - \sin\theta\, \delta_{ix})f(r) + x_i\, f'(r)\right] .$$

The field strength $F_{ij} = \partial_i A_j - \partial_j A_i$ is

$$F_{ij} = \frac{g'(r)}{er^2}\left[(\cos\theta\, \delta_{jy} - \sin\theta\, \delta_{jx})\, x_i - (\cos\theta\, \delta_{iy} - \sin\theta\, \delta_{ix})\, x_j\right] .$$

We find:

$$\partial_i\varphi^\dagger\partial_i\varphi = v^2\left(\frac{f(r)^2}{r^2} + [f'(r)]^2\right)$$

$$A_i\varphi^\dagger\partial_i\varphi - h.c. = 2i\,\frac{v^2}{er^2}\, f(r)^2 g(r)$$

$$A_i A_i\varphi^\dagger\varphi = \left(\frac{v}{er}f(r)g(r)\right)^2$$

$$(\varphi^\dagger\varphi - v^2)^2 = v^4\left(f(r)^2 - 1\right)^2$$

$$\frac{1}{4}F_{ij}F_{ij} = \frac{1}{2}\left(\frac{g'(r)}{er}\right)^2$$

Therefore, the mass of the soliton is

$$M = \int d^2x \frac{v^2}{r^2} \left\{ [g(r) - 1]^2 f(r)^2 + r^2 [f'(r)]^2 + \lambda v^2 r^2 [f(r)^2 - 1]^2 + \frac{1}{2} \left(\frac{g'(r)}{ev} \right)^2 \right\}$$

Since $[v^2] = +1$, each term in the square brackets must have mass dimension $+2$. The reader may verify using the formulas collected previously that this is true. Note that the combination $\xi \equiv evr$ is dimensionless.

Let us now switch to dimensionless variables. Define the functions of dimensionless variables $F(\xi) \equiv f(r)$ and $G(\xi) \equiv g(r)$, so that $f'(r) = ev\, F'(\xi)$ and $g'(r) = ev\, G'(\xi)$. Also define the dimensionless coupling $\beta \equiv \lambda/e^2$. In terms of these, the mass of the vortex is given by

$$M = 2\pi v^2 \int_{eva}^{evR} d\xi \frac{1}{\xi} \left\{ [G(\xi) - 1]^2 F(\xi)^2 + \xi^2 [F'(\xi)]^2 + \beta\, \xi^2 [F(\xi) - 1]^2 + \frac{1}{2} [G'(\xi)]^2 \right\}$$

where a is the size of the vortex and R is the size of the spatial region. (We are using the notation of the discussion about the Kosterlitz-Thouless phase transition on p. 310.)

We now vary with respect to the functions $F(\xi)$ and $G(\xi)$ to find the equations of motion. Minimizing with respect to $F(\xi)$ gives

$$\frac{d}{d\xi} \left(\xi \frac{d}{d\xi} F(\xi) \right) = \beta \xi [F(\xi) - 1] + \frac{1}{\xi} [G(\xi) - 1]^2 F(\xi)$$

and minimizing with respect to $G(\xi)$ gives

$$\xi \frac{d}{d\xi} \left(\frac{1}{\xi} \frac{d}{d\xi} G(\xi) \right) = 2[G(\xi) - 1] F(\xi)^2 \ .$$

Recall that the boundary conditions on $F(\xi)$ and $G(\xi)$ are that they go to 1 for $\xi \to \infty$. With this in mind, consider the ansatze

$$F(\xi) = 1 - e^{-\alpha_F \xi} \quad \text{and} \quad G(\xi) = 1 - e^{-\alpha_G \xi} \ .$$

Since $\xi = evr$, the quantities $\ell_\varphi \equiv 1/(ev\alpha_F)$ and $\ell_A \equiv 1/(ev\alpha_G)$ characterize the length scales over which the fields φ and A, respectively, vary. With the above forms, the equations of motion become

$$\alpha_F(1 - \xi\alpha_F) = -\beta\xi + \frac{1}{\xi} e^{-2\alpha_G\xi}[e^{+\alpha_F\xi} - 1] \overset{\xi \to \infty}{\longrightarrow} \alpha_F^2 = \beta \quad \text{(provided that } \alpha_F < 2\alpha_G\text{)}$$

$$(\alpha_G - \frac{1}{\xi})\alpha_G = 2[1 - e^{-\alpha_F\xi}]^2 \overset{\xi \to \infty}{\longrightarrow} \alpha_G^2 = 2$$

Thus, recalling that $\beta = \lambda/e^2$, we obtain the lengths

$$\ell_\varphi = \frac{1}{\sqrt{\lambda}\, v} \quad \text{and} \quad \ell_A = \frac{1}{\sqrt{2}\, ev} \ .$$

160

Note that the gauge coupling e has dropped out of the length scale of φ, and that the length scale of A does not depend on the quartic coupling λ. We could have arrived at this conclusion simply on the basis of dimensional analysis: both length scales are determined by $\sim 1/v$, but v has mass dimension $+1/2$ and so another power of mass dimension $+1/2$ is needed for each field. Since λ has mass dimension $+1$ and is associated with φ, we guess $\ell_\varphi \sim 1/(\sqrt{\lambda} v)$. Since e has mass dimension $+1/2$ and is associated with A, we guess $\ell_A \sim 1/(ev)$.

Note also that the condition $\alpha_F < 2\alpha_G$ implies $\lambda < 8e^2$ or $\lambda < 32\pi\alpha$, where $\alpha \equiv e^2/(4\pi)$ as usual. This implies that the configuration makes sense only if the quartic coupling for the scalar field is less than $\sim 100\alpha$.

As for the mass of the vortex, we take the above forms for $F(\xi)$ and $G(\xi)$ with $\alpha_F = \sqrt{\beta}$ and $\alpha_G = \sqrt{2}$, and perform the integral. As discussed on p. 310, we need to cut off the integral at the low end by the size a of the vortex. On the other hand, the boundary conditions discussed previously allow us to take $R \to \infty$. Asking Mathematica to perform the integral, we obtain:

$$M = 2\pi v^2 \left[\ln\left(\frac{1}{eva}\right) + \ln\left(\frac{(\sqrt{\beta} + 2\sqrt{2})^2}{16(\sqrt{\beta} + \sqrt{2})}\right) - \gamma + \frac{1}{2} + O(eva) \right]$$

where $\gamma \approx 0.577$. Perhaps a more useful quantity to compute is the mass difference between a vortex of size a' and one of size a:

$$M(a') - M(a) = 2\pi v^2 \ln\left(\frac{a}{a'}\right)$$

which shows the amount of energy required to increase the size of the vortex from a size a to a size $a' > a$. In any case, the mass of the vortex is estimated to be of order $M \sim 2\pi v^2$.

Recall the remark on p. 309 that the mass of the topological monopole in $(3 + 1)$ dimensions comes out to be of order $M \sim M_W/\alpha$, where $\alpha \equiv e^2/(4\pi)$. In our case, we have $M_W \equiv \ell_A^{-1} \sim ev \implies v^2 \sim M_W^2/e^2$ and thus in $(2 + 1)$ dimensions we find $M \sim M_W^2/e^2$, which is dimensionally correct since e now has mass dimension $+\frac{1}{2}$.

3. Consider the vortex configuration in which $\varphi(r \to \infty) \to v e^{i\nu\theta}$ with ν an integer. Calculate the magnetic flux. Show that the magnetic flux coming out of an antivortex (for which $\nu = -1$) is opposite to the magnetic flux coming out of a vortex.

Solution:

Consider the field configuration at spatial infinity

$$\varphi_\infty = v \, e^{i n \theta}$$

where n is an integer. The gauge field at infinity is

$$A_{\infty i} = \left(\frac{-i}{e}\right) \frac{\varphi_\infty^\dagger \partial_i \varphi_\infty}{|\varphi_\infty|^2} = \left(\frac{-i}{e}\right) \frac{v(v\, in\partial_i\theta)}{v^2} = \frac{n}{e}\partial_i\theta$$

Everything proceeds precisely as in the text, except multiplied by the integer n. The flux $\Phi \equiv \int d^2x\, F_{12} = \oint dx^i A_i$ is

$$\Phi = n\frac{2\pi}{e} = n\Phi_0$$

where $\Phi_0 \equiv \frac{2\pi}{e}$ is the fundamental unit of flux in equation (2) on p. 307. As discussed, a vortex corresponds to $n = +1$, while an antivortex corresponds to $n = -1$. Immediately we have

$$\Phi_{\text{vortex}} = \Phi_0 = -\Phi_{\text{antivortex}} \ .$$

6. Display explicitly the map $S^2 \to S^2$, which wraps one sphere around the other twice. Verify that this map corresponds to a magnetic monopole with magnetic charge 2.

Solution:

An arbitrary element g of $SO(3)$ can be parametrized as $g(\vec{x}) = e^{\vec{x}\cdot\vec{T}}$, where the matrices

$$T^1 \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \ , \quad T^2 \equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \ , \quad T^3 \equiv \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

satisfy $[T^a, T^b] = \varepsilon^{abc}T^c$ (with $\varepsilon^{123} \equiv +1$) and thereby generate $SO(3)$, and the coordinates

$$x^1 = \sin\theta\cos(n\phi) \ , \quad x^2 = \sin\theta\sin(n\phi) \ , \quad x^3 = \cos\theta$$

satisfy $\vec{x}\cdot\vec{x} = 1$ and thereby manifestly parameterize the surface of a two-sphere. The number n is an integer whose significance we will derive shortly. The function $g(\vec{x})$ thereby constitutes a map $S^2 \to S^2$.

We now compute the winding number of this configuration, namely the quantity

$$\nu_g \equiv -\frac{1}{8\pi}\int_{\mathbb{R}^3} \mathrm{tr}[(g^{-1}dg)^3]$$

162

where we are using forms notation. Using the notation $g = e^{\vec{x} \cdot \vec{T}}$, we have $g^{-1} dg = d\vec{x} \cdot \vec{T}$. We therefore compute the integral:

$$
\begin{aligned}
\int_{\mathbb{R}^3} tr[(g^{-1}dg)^3] &= \int_{R^3} dx^a dx^b dx^c \, \mathrm{tr}(T^a T^b T^c) \\
&= \int_{\mathbb{R}^3} dx^a dx^b dx^c \, \tfrac{1}{2}\mathrm{tr}([T^a, T^b]T^c) \\
&= \int_{\mathbb{R}^3} dx^a dx^b dx^c \, \tfrac{1}{2}\varepsilon^{abd}\mathrm{tr}(T^d T^c) \\
&= \int_{\mathbb{R}^3} dx^a dx^b dx^c \, \tfrac{1}{2}\varepsilon^{abd}(-2\delta^{dc}) \quad \leftarrow \text{ [check explicitly from the above } T^a] \\
&= -\int_{\mathbb{R}^3} dx^a dx^b dx^c \varepsilon^{abc} \\
&= -\int_{\mathbb{R}^3} d(x^a dx^b dx^c)\varepsilon^{abc} \quad \leftarrow \text{ (since } d^2 = 0) \\
&= -\int_{S^2} x^a dx^b dx^c \varepsilon^{abc} \quad \leftarrow \text{ (Stokes' theorem)} \\
&= -\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \; \varepsilon^{AB} x^a \partial_A x^b \partial_B x^c \varepsilon^{abc} \quad \leftarrow \text{ [coordinate basis } (\theta, \phi)]
\end{aligned}
$$

Therefore we have

$$
\nu_g = \frac{1}{8\pi} \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \; \varepsilon^{abc} \varepsilon^{AB} x^a \partial_A x^b \partial_B x^c \, .
$$

Using our expressions for x^a given above, we compute the partial derivatives:

$$
\begin{aligned}
\partial_\theta x^1 &= \cos\theta \cos(n\phi) \, , \quad \partial_\theta x^2 = \cos\theta \sin(n\phi) \, , \quad \partial_\theta x^3 = -\sin\theta \\
\partial_\phi x^1 &= -n \sin\theta \sin(n\phi) \, , \quad \partial_\phi x^2 = +n \sin\theta \cos(n\phi) \, , \quad \partial_\phi x^3 = 0 \, .
\end{aligned}
$$

Next we expand out the epsilon tensor ε^{abc} corresponding to the $SO(3)$ algebra directions:

$$
\begin{aligned}
\varepsilon^{abc} x^a \partial_A x^b \partial_B x^c = {}& x^1 \partial_A x^2 \partial_B x^3 + x^3 \partial_A x^1 \partial_B x^2 + x^2 \partial_A x^3 \partial_B x^1 \\
& - x^3 \partial_A x^2 \partial_B x^1 - x^1 \partial_A x^3 \partial_B x^2 - x^2 \partial_A x^1 \partial_B x^3
\end{aligned}
$$

Now we compute:

$$\varepsilon^{abc}x^a\partial_\theta x^b\partial_\phi x^c$$
$$= 0 + [\cos\theta][\cos\theta\cos(n\phi)][n\sin\theta\cos(n\phi)] + [\sin\theta\sin(n\phi)][-\sin\theta][-n\sin\theta\sin(n\phi)]$$
$$- [\cos\theta][\cos\theta\sin(n\phi)][-n\sin\theta\sin(n\phi)] - [\sin\theta\cos(n\phi)][-\sin\theta][n\sin\theta\cos(n\phi)] - 0$$

$$= n[\cos^2\theta\sin\theta\cos^2(n\phi) + \sin^2\theta\sin\theta\sin^2(n\phi)$$
$$+ \cos^2\theta\sin\theta\sin^2(n\phi) + \sin^2\theta\sin\theta\cos^2(n\phi)]$$

$$= n[\cos^2\theta\sin\theta + \sin^2\theta\sin\theta]$$

$$= n\sin\theta \ .$$

The other term contributes equally, so finally we have:

$$\nu_g = \frac{1}{4\pi}\int_0^{2\pi}d\phi\int_0^\pi d\theta\ n\sin\theta = n \ .$$

Thus the integer n is itself the winding number for the gauge field configuration. If we desire the map that wraps one sphere around the other twice, then we choose $n = 2$ in the definition of the $SO(3)$ algebra coordinates x^a.

Next we have to compute the magnetic charge of the monopole described by this field configuration. Equation (7) on p. 308 provides us with the definition of the gauge-invariant electromagnetic field:

$$\mathcal{F}_{ij} \equiv \vec{F}_{ij}\cdot\vec{x} - \frac{1}{e}\varepsilon^{abc}x^a(D_ix)^b(D_jx)^c$$

where $(D_ix)^a = \partial_ix^a + e\,\varepsilon^{abc}A_i^bx^c$ and $A_i^b = \partial_ix^b$ as before, and we have used $x^ax^a = 1$. (We are labeling the field variables by x^a – as we have been throughout this problem – to emphasize that the parameters on the group manifold should be thought of as coordinates like any other. We use i, j, k to denote spatial $\mathbb{R}^3$, and A, B to denote the S^2 of spatial infinity.)

The gauge-*covariant* field strength is $F_{ij}^a = \partial_iA_j^a - \partial_jA^a + e\,\varepsilon^{abc}A_i^bA_j^c = e\,\varepsilon^{abc}\partial_ix^b\partial_jx^c$, so $\vec{F}_{ij}\cdot\vec{x} = e\,\varepsilon^{abc}x^a\partial_ix^b\partial_jx^c$.

Next we have to contract two covariant derivatives with an epsilon. This will require using $\varepsilon^{abc}\varepsilon^{ade} = \delta^{bd}\delta^{ce} - \delta^{be}\delta^{cd}$, which follows from $SO(3)$ invariance and checking with $b = d = 2$ and $c = e = 3$. We compute:

$$x^a \varepsilon^{abc}(D_i x)^b (D_j x)^c = x^a [\varepsilon^{abc} \partial_i x^b \partial_j x^c + e \partial_i x^b (\varepsilon^{abc} \varepsilon^{cpq}) \partial_j x^p x^q$$
$$+ e(\varepsilon^{abc} \varepsilon^{bef}) \partial_i x^e x^f \partial_j x^c + e^2 (\varepsilon^{abc} \varepsilon^{bef} \varepsilon^{cpq}) \partial_i x^e x^f \partial_j x^p x^q]$$

$$= x^a [\varepsilon^{abc} \partial_i x^b \partial_j x^c + e \partial_i x^b (\delta^{ap} \delta^{bq} - \delta^{aq} \delta^{pb}) \partial_j x^p x^q$$
$$+ e(-\delta^{ae} \delta^{cf} + \delta^{af} \delta^{ec}) \partial_i x^e x^f \partial_j x^c + e^2 (\delta^{ap} \varepsilon^{qef} - \delta^{aq} \varepsilon^{pef}) \partial_i x^e x^f \partial_j x^p x^q]$$

$$= \varepsilon^{abc} x^a \partial_i x^b \partial_j x^c + e \underbrace{[(\partial_i x^b \partial_j x^a x^b x^a - \partial_i x^b \partial_j x^b) + (-x^a \partial_i x^a x^b \partial_j x^b + \partial_i x^a \partial_j x^a)]}_{\text{cancel}}$$
$$+ e^2 [(\underbrace{\varepsilon^{qef} x^f x^q}_{=0} \partial_i x^e \partial_j x^a x^a) - (\underbrace{\varepsilon^{pef} \partial_i x^e x^f \partial_j x^p}_{= -\varepsilon^{abc} x^a \partial_i x^b \partial_j x^c})]$$

$$= (1 + e^2) \varepsilon^{abc} x^a \partial_i x^b \partial_j x^c$$

Therefore, the gauge-*invariant* field strength is

$$\mathcal{F}_{ij} = e \, \varepsilon^{abc} x^a \partial_i x^b \partial_j x^c - \frac{1}{e}[(1 + e^2) \varepsilon^{abc} x^a \partial_i x^b \partial_j x^c]$$
$$= -\frac{1}{e} \varepsilon^{abc} x^a \partial_i x^b \partial_j x^c \ .$$

The magnetic field is

$$B_i \equiv \frac{1}{2} \varepsilon_{ijk} \mathcal{F}_{jk} = -\frac{1}{2e} \varepsilon_{ijk} \varepsilon^{abc} x^a \partial_i x^b \partial_j x^c \ .$$

The magnetic flux through a 2-sphere at infinity is

$$g = -\frac{1}{2e} \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \, \varepsilon^{abc} \varepsilon^{AB} x^a \partial_A x^b \partial_B x^c$$
$$= -\frac{1}{2e}(8\pi n)$$
$$= -\frac{4\pi}{e} n \ .$$

As explained in the book's solution for V.7.5, the fundamental unit of charge is actually $\frac{1}{2}e$, so that we have Dirac's quantization condition

$$g = -\frac{2\pi}{(e/2)} n$$

as expected. As in V.7.5, the sign just corresponds to which we call the monopole and which one we call the antimonopole.

9. Discuss the dyon solution. Work it out in the BPS limit. [B. Julia and A. Zee, *Phys. Rev. D11*: 2227, 1975.]

Solution:

We follow M. K. Prasad and C. M. Sommerfield, "Exact classical solution for the 't Hooft monopole and the Julia-Zee dyon," Phys. Rev. Lett. Vol. 35, No. 12, 22 Sep 1975.

The Lagrangian density is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} + \frac{1}{2}(D_\mu\phi)^a(D^\mu\phi)^a - V(\phi)$$
$$= -\frac{1}{4}F_{ij}^a F_{ij}^a + \frac{1}{2}F_{0i}^a F_{0i}^a - \frac{1}{2}(D_i\phi)^a(D_i\phi)^a + \frac{1}{2}(D_0\phi)^a(D_0\phi)^a - V(\phi)$$

with potential $V(\phi) = -\frac{1}{2}\mu^2(\phi^a\phi^a) + \frac{1}{4}\lambda(\phi^a\phi^a)^2$ ($\mu^2 > 0$) and covariant derivative $(D_\mu\phi)^a = \partial_\mu\phi^a + e\varepsilon^{abc}A_\mu^b\phi^c$. The field strength is $F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + \varepsilon^{abc}A_\mu^b A_\nu^c$. The equations of motion are

$$(D^\mu F_{\mu\nu})^a + e\varepsilon^{abc}\phi^b(D_\nu\phi)^c = 0$$
$$(D^\mu D_\mu\phi)^a - \mu^2\phi^a + \lambda(\phi^b\phi^b)\phi^a = 0 \ .$$

There is also a constraint $(D^\mu F_{\mu 0})^a + e\varepsilon^{abc}\phi^b(D_0\phi)^c = 0$.

Choose the ansatz

$$\phi^a = \frac{x^a}{er^2}H(r) \ , \quad A_i^a = \varepsilon^{aij}\frac{x^j}{er^2}[1 - K(r)] \ , \quad A_0^a = \frac{x^a}{er^2}J(r)$$

and plug into the equations of motion (including the constraint) to obtain:

$$r^2 K'' = K(K^2 - 1) + K(H^2 - J^2)$$
$$r^2 J'' = 2JK^2$$
$$r^2 H'' = 2HK^2 + \frac{\lambda}{e^2}(H^2 - C^2 r^2)H$$

where $C \equiv \mu e/\lambda^{1/2}$. The boundary conditions are:

$$H(0) = J(0) = 0 \ , \quad K(0) = 1$$
$$K(r \to \infty) = 0$$
$$\lim_{r\to\infty} \frac{J(r)}{r} = M$$
$$\lim_{r\to\infty} \frac{H(r)}{r} = C\cosh\gamma$$

where M is the characteristic inverse length of A_0^a, and γ is an arbitrary constant.

The equations of motion admit the solution

$$K = \frac{Cr}{\sinh(Cr)}$$

$$J = \sinh \gamma \left[Cr \coth(Cr) - 1 \right]$$

$$H = \cosh \gamma \left[Cr \coth(Cr) - 1 \right] .$$

Given the electromagnetic field (see equation (7) on p. 308 of the book)

$$\mathcal{F}_{\mu\nu} = \frac{1}{|\phi|} \phi^a F^a_{\mu\nu} - \frac{1}{e|\phi|^3} \varepsilon^{abc} \phi^a (D_\mu \phi)^b (D_\nu \phi)^c$$

we find the electric field $\mathcal{E}_i \equiv -\mathcal{F}_{0i}$ to be

$$\mathcal{E}_i = \frac{x_i}{r} \frac{d}{dr} \left(\frac{J(r)}{er} \right) = \frac{x_i}{r} \sinh \gamma \left(\frac{1 - K^2}{er^2} \right) .$$

This can be used to find the electric charge:

$$Q = \int d^3x \, \partial_i \mathcal{E}_i = \frac{8\pi}{e} \int_0^\infty dr \, \frac{JK^2}{r} = \frac{4\pi}{e} \sinh \gamma .$$

The case $\gamma = 0$ corresponds to the monopole solution discussed in the text. See both references for further discussion.

10. Verify explicitly that the magnetic monopole is rotation invariant in spite of appearances. By this is meant that all physical gauge invariant quantities such as $\vec{B}$ are covariant under rotation. Gauge variant quantities such as A_i^b can and do vary under rotation. Write down the generators of rotation.

Solution:

First a small but important typographical error: All physical gauge invariant quantities are *invariant* under rotation, not covariant. Covariant means transforms as a vector, while invariant means does not transform. Gauge transformations (rotations within the gauge group) are conceptually nothing but changing coordinates (in field space), and physics must be invariant under a change of coordinates.

The electromagnetic field from the magnetic monopole is

$$\mathcal{F}_{\mu\nu} = \frac{F^a_{\mu\nu} \varphi^a}{|\varphi|} - \frac{\varepsilon^{abc} \varphi^a (D_\mu \varphi)^b (D_\nu \varphi)^c}{e|\varphi|^3}$$

where $|\varphi| = \sqrt{\varphi^a \varphi^a}$, $(D_i \varphi)^a = \partial_i \varphi^a + e\varepsilon^{abc} A_i^b \varphi^c$ and

$$\varphi^a(r) = \frac{v}{r} \delta^{ai} x^i f(r) , \quad A_i^a(r) = \frac{1}{er^2} \varepsilon^{aij} x^j g(r) , \quad A_0^a = 0$$

167

with $f(r)$ and $g(r)$ yet unspecified functions of r. Since we are concerned only with the transformation properties of the various terms under $SO(3)$ gauge transformations, we do not need to specify the functions $f(r)$ and $g(r)$ other than saying that they are scalars under gauge transformations, as indicated by their lack of $SO(3)$ indices.

Since the symbols δ^{ab} and ε^{abc} are invariant under $SO(3)$, the electromagnetic field $\mathcal{F}_{\mu\nu}$ is manifestly gauge invariant. We will now verify this explicitly using infinitesimal $SO(3)$ transformations.

An arbitrary infinitesimal $SO(3)$ transformation can be parametrized as $U(n, \theta) = I + \theta n^a T^a$, where $n^a n^a = 1$ and the generators T^a are

$$T^1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad T^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad T^3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and satisfy the commutation relations $[T^a, T^b] = \varepsilon^{abc} T^c$, where $\varepsilon^{123} \equiv +1$. (By the way, this immediately shows that ε^{abc} is invariant under $SO(3)$, since the structure constants of a Lie algebra are invariant symbols of the corresponding group.)

To show explicitly that $\mathcal{F}_{\mu\nu}$ is invariant under $SO(3)$, we need to show explicitly that δ^{ab} and ε^{abc} are invariant under simultaneous $SO(3)$ transformations on all of their indices, since the first term of $\mathcal{F}_{\mu\nu}$ is constructed from δ^{ab} and the second term is constructed from ε^{abc} (and the magnitude $|\varphi|$ is also constructed from δ^{ab}).

Note that the indices a, b, c label the adjoint representation of $SO(3)$.

First, transform δ^{ab}. Under an infinitesimal $SO(3)$ transformation, this symbol transforms as

$$\begin{aligned}
\delta^{ab} \to U^{ac} U^{bd} \delta^{cd} &= [I + i\theta n^e T^e]^{ac}[I + i\theta n^f T^f]^{bd} \delta^{cd} \\
&= \delta^{ab} + i\theta n^e (T^e)^{ac} \delta^{cb} + i\theta n^f (T^f)^{bd} \delta^{ad} + O(\theta^2) \\
&= \delta^{ab} + i\theta n^e (T^e)^{ab} + i\theta n^f (T^f)^{ba} \\
&= \delta^{ab} \checkmark
\end{aligned}$$

since $(T^a)^{bc} = -(T^a)^{cb}$, as can be seen above by the explicit construction of the matrices $\{T^a\}_{a=1}^3$.

Now transform ε^{abc}:

$$\begin{aligned}
\varepsilon^{abc} \to U^{ad} U^{be} U^{cf} \varepsilon^{def} &= [I + i\theta \vec{n} \cdot \vec{T}]^{ad}[I + i\theta \vec{n} \cdot \vec{T}]^{be}[I + i\theta \vec{n} \cdot \vec{T}]^{cf} \varepsilon^{def} \\
&= \varepsilon^{abc} + i\theta \vec{n} \cdot \left(\vec{T}^{ad} \varepsilon^{dbc} + \vec{T}^{be} \varepsilon^{aec} + \vec{T}^{cf} \varepsilon^{abf} \right) + O(\theta^2) \\
&= \varepsilon^{abc} + i\theta \vec{n} \cdot \left(\vec{T}^{ad} \varepsilon^{dbc} + \vec{T}^{bd} \varepsilon^{adc} + \vec{T}^{cd} \varepsilon^{abd} \right)
\end{aligned}$$

Specialize to the case $(abc) = (123)$, since all elements of ε^{abc} are defined in terms of ε^{123}. We have

$$\varepsilon^{123} \rightarrow \varepsilon^{123} + i\theta\vec{n} \cdot \left(\vec{T}^{1d}\varepsilon^{d23} + \vec{T}^{2d}\varepsilon^{1d3} + \vec{T}^{3d}\varepsilon^{12d} \right)$$

$$= \varepsilon^{123} + i\theta\vec{n} \cdot \left(\vec{T}^{11}\varepsilon^{123} + \vec{T}^{22}\varepsilon^{123} + \vec{T}^{33}\varepsilon^{123} \right)$$

$$= \varepsilon^{123} \checkmark$$

since all diagonal entries of the T^a are equal to zero. This completes the calculation.

VI Field Theory and Condensed Matter

VI.1 Fractional Statistics, Chern-Simons Term, and Topological Field Theory

1. In a nonrelativistic theory you might think that there are two separate Chern-Simons terms, $\varepsilon_{ij}a_i\partial_0 a_j$ and $\varepsilon_{ij}a_0\partial_i a_j$. Show that gauge invariance forces the two terms to combine into a single Chern-Simons term $\varepsilon^{\mu\nu\lambda}a_\mu\partial_\nu a_\lambda$. For the Chern-Simons term, gauge invariance implies Lorentz invariance. In contrast, the Maxwell term would in general be nonrelativistic, consisting of two terms, f_{0i}^2 and f_{ij}^2, with an arbitrary relative coefficient between them (with $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$ as usual).

Solution:

The Lagrangian is

$$\mathcal{L} = c_1 \varepsilon^{ij} a_i \partial_0 a_j + c_2 \varepsilon^{ij} a_0 \partial_i a_j \ .$$

A gauge transformation $a_\mu \rightarrow a_\mu + \partial_\mu \lambda$ implies $\mathcal{L} \rightarrow \mathcal{L} + (2c_1 + c_2)\varepsilon^{ij}\partial_0 \lambda \partial_i a_j \implies c_2 = -2c_1$. Up to total derivatives, $\varepsilon^{\mu\nu\rho}a_\mu \partial_\nu a_\rho = -\varepsilon^{0ij}(a_i\partial_0 a_j - 2a_0\partial_i a_j)$. Therefore $\mathcal{L} = -c_1 \varepsilon^{\mu\nu\rho} a_\mu \partial_\nu a_\rho$.

2. By thinking about mass dimension, convince yourself that the Chern-Simons term dominates the Maxwell term at long distances. This is one reason that relativistic field theorists find anyon fluids so appealing. As long as they are interested only in long distance physics they can ignore the Maxwell term and play with a relativistic theory (see exercise VI.1.1). Note that this picks out (2+1)-dimensional spacetime as special. In (3+1)-dimensional spacetime the generalization of the Chern-Simons term $\varepsilon^{\mu\nu\lambda\sigma}f_{\mu\nu}f_{\lambda\sigma}$ has the same mass dimension as the Maxwell term f^2. In (4+1)-dimensional space the term $\varepsilon^{\rho\mu\nu\lambda\sigma}a_\rho f_{\mu\nu}f_{\lambda\sigma}$ is less important at long distances than the Maxwell term f^2.

Solution:

There are two standard ways to assign mass dimensions to gauge fields, which of course yield equivalent results. If we normalize the gauge field by $\mathcal{L} = -\frac{1}{4g^2}f^2$ with the covariant derivative $D_\mu \psi = (\partial_\mu - ia_\mu)\psi$ (for some matter field ψ), then $[a] = [\partial] = +1$ for any choice of the spacetime dimension d. The Chern-Simons term $\varepsilon^{\mu\nu\rho}a_\mu \partial_\nu a_\rho$ therefore has mass dimension $[a\partial a] = 2[a] + [\partial] = 3$. Meanwhile, the field strength $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$ has dimension $[f] = [\partial a] = 2$, so the Maxwell term f^2 has dimension 4. Therefore $[a\partial a] < [f^2]$.

Instead suppose we normalize the gauge field by $\mathcal{L} = -\frac{1}{4}f^2$ with the covariant derivative $D_\mu \psi = (\partial_\mu - iga_\mu)\psi$. Since the Lagrangian always has dimension $[\mathcal{L}] = d$ in d spacetime dimensions to make the action $S = \int d^d x\, \mathcal{L}$ dimensionless, we have $[f^2] = 2[\partial a] = 2(1+[a]) = d \implies [a] = (d-2)/2$. The Chern-Simons term has dimension $[a\partial a] = 2[a] + 1 = d - 2 + 1 = d - 1$. Since the Maxwell term f^2 has the same mass dimension as that of the Lagrangian $\mathcal{L}$, namely d, we see again that $[a\partial a] < [f^2]$.

3. There is a generalization of the Chern-Simons term to higher dimensional spacetime different from that given in exercise IV.1.2. We can introduce a p-form gauge potential (see chapter IV.4). Write the generalized Chern-Simons term in $(2p + 1)$-dimensional spacetime and discuss the resulting theory.

Solution:

First a typo: The problem meant to contrast this with the previous exercise, VI.1.2.

In $(2p + 1)$-dimensional spacetime, the Chern-Simons term is a $(2p + 1)$-form. If we introduce a p-form gauge potential A, then $F \equiv dA$ is a $(p + 1)$-form. The quantity AF is a $(p + p + 1) = (2p + 1)$-form, so the Chern-Simons Lagrangian is $\mathcal{L}_{\mathrm{CS}} = \gamma\, AF$, where γ is a coupling constant.

4. Consider $\mathcal{L} = \gamma a\varepsilon\partial a - (1/4g^2)f^2$. Calculate the propagator and show that the gauge boson is massive. Some physicists puzzled by fractional statistics have reasoned that since in the presence of the Maxwell term the gauge boson is massive and hence short ranged, it can't possibly generate fractional statistics, which is manifestly an infinite ranged interaction. (No matter how far apart the two particles we are interchanging are, the wave function still acquires a phase.) The resolution is that the information is in fact propagated over an infinite range by a $q = 0$ pole associated with a gauge degree of freedom. This apparent paradox is intimately connected with the puzzlement many physicists felt when they first heard of the Aharonov-Bohm effect. How can a particle in a region with no magnetic field whatsoever and arbitrarily far form the magnetic flux know about the existence of the magnetic flux?

Solution:

The Lagrangian is

$$\mathcal{L} = -\frac{1}{4g^2} f_{\mu\nu} f^{\mu\nu} + \gamma \varepsilon^{\mu\nu\rho} a_\mu \partial_\nu a_\rho$$

Since $f_{\mu\nu} f^{\mu\nu} = 2a^\mu (-\eta_{\mu\nu} \partial^2 + \partial_\mu \partial_\nu) a^\nu$, we have

$$\mathcal{L} = \frac{1}{2} a^\mu \left[\frac{1}{g^2} (\eta_{\mu\nu} \partial^2 - \partial_\mu \partial_\nu) + 2\gamma \varepsilon_{\mu\lambda\nu} \partial^\lambda \right] a^\nu$$

The propagator G is the inverse of the thing in brackets, so

$$\left[\frac{1}{g^2} (\eta_{\mu\nu} \partial^2 - \partial_\mu \partial_\nu) + 2\gamma \varepsilon_{\mu\lambda\nu} \partial^\lambda \right] G^{\nu\rho}(x) = \delta_\mu{}^\rho \delta^3(x)$$

Multiplying by e^{-ikx}, integrating $\int d^3x$ and moving derivatives around via integration by parts gives

$$\left[-\frac{1}{g^2} (\eta_{\mu\nu} k^2 - k_\mu k_\nu) + 2i\gamma \varepsilon_{\mu\lambda\nu} k^\lambda \right] \tilde{G}^{\nu\rho}(k) = \delta_\mu{}^\rho$$

If we try to solve for $\tilde{G}^{\nu\rho}$ now we will run into trouble. To see why, use Lorentz invariance to write

$$\tilde{G}^{\nu\rho}(k) = A(k^2)\eta^{\nu\rho} + B(k^2) k^\nu k^\rho + C(k^2) \varepsilon^{\nu\rho\sigma} k_\sigma \ .$$

Putting this into the equation for $\tilde{G}^{\nu\rho}$ gives

$$\left(\frac{A}{g^2} - 2i\gamma C \right) \left(k^2 \delta_\mu^\rho - k^\rho k_\mu \right) + \left(2i\gamma A + \frac{C}{g^2} k^2 \right) \varepsilon_\mu{}^{\rho\lambda} k_\lambda = -\delta_\mu^\rho$$

which has no solution. (Note that B dropped out.) This is just the usual gauge fixing problem, which we can resolve by adding a term $\frac{1}{2\xi}\eta_{\mu\nu}$ to the equation for $\tilde{G}^{\nu\rho}$:

$$\left[-\frac{1}{g^2} (\eta_{\mu\nu} k^2 - k_\mu k_\nu) + 2i\gamma \varepsilon_{\mu\lambda\nu} k^\lambda + \frac{1}{2\xi}\eta_{\mu\nu} \right] \tilde{G}^{\nu\rho}(k) = \delta_\mu{}^\rho$$

Now plugging in the form $\tilde{G} = A\eta + Bkk + C\varepsilon k$ results in the conditions

$$\left(k^2 - \frac{g^2}{2\xi} \right) \frac{A}{g^2} - 2i\gamma C k^2 = -1 \ , \quad \frac{A}{g^2} + \frac{B}{2\xi} - 2i\gamma C = 0 \ , \quad 2i\gamma A + \frac{C}{g^2} \left(k^2 - \frac{g^2}{2\xi} \right) = 0 \ .$$

Solving these gives

$$A = \frac{-(k^2 - \frac{g^2}{2\xi})g^2}{(k^2 - \frac{g^2}{2\xi})^2 - (2\gamma g^2)^2 k^2} \quad \overset{\xi \to \infty}{\to} \quad \frac{-g^2}{k^2 - (2\gamma g^2)^2}$$

$$B = \frac{-g^2}{(k^2 - \frac{g^2}{2\xi})^2 - (2\gamma g^2)^2 k^2} \left[1 - \frac{2\xi}{g^2} \left(k^2 - (2\gamma g^2)^2 \right) \right] \quad \overset{\xi \to \infty}{\to} \quad \frac{2\xi}{k^2}$$

$$C = \frac{2i\gamma g^4}{(k^2 - \frac{g^2}{2\xi})^2 - (2\gamma g^2)^2 k^2} \quad \overset{\xi \to \infty}{\to} \quad \frac{2i\gamma g^4}{k^2(k^2 - (2\gamma g^2)^2)} \ .$$

171

The propagator for the gauge boson has a simple pole at $k^2 = (2\gamma g^2)^2$ which persists in the limit $\xi \to \infty$. Therefore, the gauge boson has a mass $m = 2\gamma g^2$. The propagator also has a simple pole at $k^2 = 0$; notice that the terms involving the pole at $k^2 = m^2$ are independent of the gauge parameter ξ when taking $\xi \to \infty$, while the pole at $k^2 = 0$ comes with a factor of ξ in the numerator in the coefficient B. This is what is meant by the $k^2 = 0$ pole being associated with a gauge degree of freedom. Even though the physical gauge boson has mass $m = 2\gamma g^2 \neq 0$, there is still a long-range interaction from the nontrivial topology of the gauge theory.

5. Show that $\theta = 1/4\gamma$. There is a somewhat tricky factor of 2. So if you are off by a factor of 2, don't despair. Try again. [X. G. Wen and A. Zee, *J. de Physique*, 50: 1623, 1989.]

Solution:

As instructed at the top of p. 318, one approach is to take the Lagrangian $\mathcal{L}_a = \gamma \varepsilon^{\mu\nu\lambda} a_\mu \partial_\nu a_\lambda + a_\mu j^\mu$ and integrate out a_μ to get the nonlocal Hopf Lagrangian. In anticipation of fixing transverse gauge $\partial_\mu a^\mu = 0$, add a term $-\frac{1}{2\xi}(\partial a)^2$ to the Lagrangian to get, after integration by parts and ignoring a total derivative, the gauge-fixed Lagrangian $\mathcal{L}_a^\xi = \frac{1}{2} a_\mu (G^{-1})^{\mu\lambda} a_\lambda$, where $(G^{-1})^{\mu\lambda} = 2\gamma \varepsilon^{\mu\nu\lambda} \partial_\nu + \frac{1}{\xi} \partial^\mu \partial^\lambda$. The effective action S_{Hopf} is computed by performing the gauge-fixed path integral over a_μ:

$$S_{\text{Hopf}} = -i \ln \int \mathcal{D}a \, e^{i \int d^3x \, \mathcal{L}_a^\xi} = \int d^3x \, d^3y \, j^\lambda(x) G_{\lambda\rho}(x-y) j^\rho(y)$$

where G is the operator inverse of G^{-1} defined by the equation

$$\left(2\gamma \varepsilon^{\mu\nu\lambda} \partial_\nu + \frac{1}{\xi} \partial^\mu \partial^\lambda\right) G_{\lambda\rho}(x) = \delta^\mu_\rho \delta^3(x) \, .$$

Multiplying by $\int d^3x \, e^{-ikx}$ gives the momentum-space equation

$$\left(+2\gamma \varepsilon^{\mu\nu\lambda} i k_\nu - \frac{1}{\xi} k^\mu k^\lambda\right) \tilde{G}_{\lambda\rho}(k) = \delta^\mu_\rho$$

where $\tilde{G}(k) \equiv \int d^3x \, e^{-ikx} G(x)$. We now need to solve for $\tilde{G}$ in Lorentz-index space. By Lorentz invariance, we can write $\tilde{G}_{\lambda\rho}(k)$ as

$$G_{\lambda\rho}(k) = A(k^2) \eta_{\lambda\rho} + B(k^2) \varepsilon_{\lambda\rho\sigma} k^\sigma + C(k^2) k_\lambda k_\rho \, .$$

Plugging this form into the above equation for $\tilde{G}(k)$ and using $\varepsilon^{\mu\nu\lambda} \varepsilon_{\lambda\rho\sigma} = \delta^\mu_\rho \delta^\nu_\sigma - \delta^\mu_\sigma \delta^\nu_\rho$ results in the three conditions:

$$2i\gamma B k^2 = 1 \, , \quad 2i\gamma A \varepsilon^{\mu\nu}_{\ \ \rho} k_\nu = 0 \, , \quad 2i\gamma B + \frac{1}{\xi} C k^2 = 0 \, .$$

Therefore $A = 0$ and
$$B = \frac{1}{2i\gamma k^2} , \quad C = -\frac{\xi}{(k^2)^2} .$$

The momentum-space propagator is therefore

$$\tilde{G}_{\lambda\rho}(k) = \frac{1}{2i\gamma k^2}\varepsilon_{\lambda\rho\sigma}k^\sigma - \frac{\xi}{(k^2)^2}k_\lambda k_\rho \overset{\xi \to 0}{\longrightarrow} \frac{1}{2i\gamma k^2}\varepsilon_{\lambda\rho\sigma}k^\sigma .$$

Now the goal is to evaluate $S_{\text{Hopf}} = \int d^3x\, d^3y\, j^\lambda(x)G_{\lambda\rho}(x-y)j^\rho(y)$ with

$$G_{\lambda\rho}(x-y) = \int \frac{d^3k}{(2\pi)^3}\frac{e^{+ik(x-y)}}{2i\gamma k^2}\varepsilon_{\lambda\rho\sigma}k^\sigma$$

and the current $j^\lambda(x) = j_1^\lambda(x) + j_2^\lambda(x)$ describing the exchange of two particles. Let particle 1 with charge q_1 be sitting at rest at the origin: $j_1^\lambda(x) = q_1\delta_0^\lambda\delta^2(\vec{x})$. We want particle 2 with charge q_2 to move in a half-circle of radius R around particle 1: $j_2^\lambda(y) = q_2 u^\lambda(y)\delta^2(\vec{y}-\vec{r}(t))$, where

$$t \equiv y^0 \in [0, \frac{\pi}{\omega}], \quad \vec{r}(t) = R(\hat{x}\cos(\omega t)+\hat{y}\sin(\omega t)), \quad u^\lambda(y) = (1, \dot{\vec{y}}).$$

The speed $\dot{\vec{y}}$ can be taken arbitrarily small to ignore the relativistic factor $(1-|\dot{\vec{y}}|^2)^{-1/2}$.

We want the cross term $S_{\text{Hopf}}^{\text{cross}} = \int d^3x\, d^3y\, j_1^\lambda(x)G_{\lambda\rho}(x-y)j_2^\rho(y)$. Plugging in gives

$$S_{\text{Hopf}}^{\text{cross}} = \int d^3x \int d^3y\, \left(q_1\delta_0^\lambda\delta^2(\vec{x})\right)\left(\int \frac{d^3k}{(2\pi)^3}\frac{e^{+ik(x-y)}}{2i\gamma k^2}\varepsilon_{\lambda\rho\sigma}k^\sigma\right)\left(q_2 u^\lambda(y)\delta^2(\vec{y}-\vec{r}(t))\right)$$

$$= q_1 q_2 \int dx^0 \int_0^{\pi/\omega} dt \int \frac{d^3k}{(2\pi)^3}\frac{e^{+ik^0(x^0-t)+i\vec{k}\cdot\vec{r}(t)}}{2i\gamma[(k^0)^2-|\vec{k}|^2]}\varepsilon_{0ij}k^j u^i(t,\vec{r}(t))$$

Since $\int_{-\infty}^\infty dx^0\, e^{+ik^0x^0} = 2\pi\delta(k^0)$, we can perform the integral over x^0 and then the integral $\int dk^0$, leaving behind the above expression with k^0 set to zero:

$$S_{\text{Hopf}}^{\text{cross}} = q_1 q_2 \int_0^{\pi/\omega} dt \int \frac{d^2k}{(2\pi)^2}\frac{e^{+i\vec{k}\cdot\vec{r}(t)}}{-2i\gamma|\vec{k}|^2}\varepsilon_{0ij}k^j\dot{r}^i(t) .$$

The arbitrary 2-momentum k^j can be written in polar coordinates as $\vec{k} = k(\hat{x}\cos\varphi+\hat{y}\sin\varphi)$, so that $\vec{k}\cdot\vec{r}(t) = kR[\cos(\omega t)\cos\varphi + \sin(\omega t)\sin\varphi] = kR\cos(\omega t - \varphi)$. Also, since $\dot{\vec{r}}(t) = \omega R(-\hat{x}\sin(\omega t)+\hat{y}\cos(\omega t))$ we have

$$\varepsilon_{0ij}\dot{r}^i(t)k^j = \omega kR[-\sin(\omega t)\sin\varphi + (-1)\cos(\omega t)\sin\varphi] = -\omega kR\cos(\omega t - \varphi)$$

where the (-1) comes from $\varepsilon_{021} = -\varepsilon_{012} = -1$. In polar coordinates, $\int d^2k = \int_0^{2\pi} d\varphi \int_0^\infty dk\, k$, so all of the factors of $k = |\vec{k}|$ cancel out in the integrand. We have

$$
\begin{aligned}
S_{\text{Hopf}}^{\text{cross}} &= \frac{q_1 q_2}{-2i\gamma}(-\omega R) \int_0^{\pi/\omega} dt \int_0^{2\pi} \frac{d\varphi}{2\pi} \int_0^\infty \frac{dk}{2\pi} e^{ikR\cos(\omega t - \varphi)} \cos(\omega t - \varphi) \\
&= \frac{q_1 q_2 \omega R}{i\gamma} \int_0^{\pi/\omega} dt \int_0^{2\pi} \frac{d\varphi}{2\pi} \frac{1}{iR} \int_0^\infty \frac{dk}{2\pi} \frac{\partial}{\partial k} e^{ikR\cos(\omega t - \varphi)} \\
&= \frac{q_1 q_2 \omega}{(-1)2\gamma} \int_0^{\pi/\omega} dt \int_0^{2\pi} \frac{d\varphi}{2\pi} \frac{1}{2\pi} \left[e^{ikR\cos(\omega t - \varphi)} \right]_{k=0}^\infty \\
&= -\frac{q_1 q_2 \omega}{4\pi\gamma} \int_0^{\pi/\omega} dt \int_0^{2\pi} \frac{d\varphi}{2\pi} \left[\underbrace{\lim_{k\to\infty} e^{ikR\cos(\omega t - \varphi)}}_{=0} - 1 \right] \\
&= -\frac{q_1 q_2 \omega}{4\pi\gamma} \left(-\frac{\pi}{\omega}\right) = +\frac{q_1 q_2}{4\gamma}
\end{aligned}
$$

Therefore the statistics parameter $\theta = S_{\text{Hopf}}$ for two particles of charge $q_1 = q_2 = +1$ is $\theta = 1/(4\gamma)$.

6. Find the nonabelian version of the Chern-Simons term ada. [Hint: As in chapter IV.6 it might be easier to use differential forms.]

Solution:

The new feature in a non-abelian theory is that $a^3 \neq 0$. The Lagrangian will be of the form $\mathcal{L} = \gamma \operatorname{tr}(ada + \beta a^3)$ where the coefficient β should be fixed by gauge invariance (up to a total derivative). For a gauge transformation $a \to a + \delta a$, we get $\delta\mathcal{L} = \operatorname{tr}[(2\, da + 3\beta\, a^2)\delta a]$, where we have dropped a total derivative. A non-abelian gauge transformation $\delta a = [\theta, a] - d\theta$ with matrix-valued infinitesimal parameter $\theta(x)$ implies $\delta\mathcal{L} = \gamma \operatorname{tr}[(2 - 3\beta)a^2\, d\theta]$, where again we have dropped total derivatives. Demanding that $\delta\mathcal{L} = 0$ implies $\beta = \frac{2}{3}$, so the non-abelian Chern-Simons Lagrangian is

$$
\mathcal{L} = \gamma \operatorname{tr}\left(ada + \tfrac{2}{3}a^3\right) .
$$

7. Using the canonical formalism of chapter I.8 show that the Chern-Simons Lagrangian leads to the Hamiltonian $H = 0$.

Solution:

The Chern-Simons Lagrangian is

$$
\mathcal{L} = \gamma \varepsilon^{\mu\nu\rho} a_\mu \partial_\nu a_\rho = 2\gamma\, a_0 B + \gamma \varepsilon^{0ij} \dot{a}_i a_j
$$

174

where in the second equality we have integrated by parts, dropped a total derivative and defined $B \equiv \frac{1}{2}\varepsilon^{0ij}F_{ij}$. Notice that a_0 has no time derivatives and is therefore constant; it acts as a Lagrange multiplier to enforce the constraint $B = 0$:

$$\frac{\partial \mathcal{L}}{\partial a_0} = 0 \implies B = 0 .$$

The Lagrangian is now

$$\mathcal{L} = \gamma \varepsilon^{0ij} \dot{a}_i a_j$$

which is linear in time derivatives. The conjugate momenta to a_i are

$$\pi^i \equiv \frac{\partial \mathcal{L}}{\partial \dot{a}_i} = \gamma \varepsilon^{0ij} a_j$$

and therefore

$$\mathcal{H} \equiv \pi^i \dot{a}_i - \mathcal{L} = 0 .$$

8. Evaluate (6).

$$\int \frac{d^3p}{(2\pi)^3} \, \text{tr} \, \left(\gamma^\nu \frac{1}{\not{p} + \not{q} - m} \gamma^\mu \frac{1}{\not{p} - m} \right) \qquad (6)$$

Solution:

First we need gamma matrix identities in 3 dimensions. As pointed out in problem II.1.12, in (2+1) dimensions we can take the gamma matrices to be $\gamma^0 = \sigma^3$, $\gamma^1 = i\sigma^2$ and $\gamma^2 = -i\sigma^1$, where σ^i are the usual 2-by-2 Pauli matrices. Let a_μ, b_μ, c_μ and d_μ be arbitrary 4-vectors. From the defining relation of the Clifford algebra $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}I$, where here I is the 2×2 identity matrix, and using Lorentz invariance we get the following gamma matrix relations:

$$\text{tr}(\not{a}\not{b}) = 2(ab) \quad \text{where} \quad (ab) \equiv a \cdot b \equiv \eta^{\mu\nu}a_\mu b_\nu$$
$$\text{tr}(\not{a}\not{b}\not{c}) = -2i\,\varepsilon^{\mu\nu\rho}a_\mu b_\nu c_\rho$$
$$\text{tr}(\not{a}\not{b}\not{c}\not{d}) = 2\left((ab)(cd) + (ad)(bc) - (ac)(bd)\right)$$

Note that unlike in four dimensions, the trace of three gamma matrices is not zero.

Next notice that this integral is linearly divergent, so we will have define it properly such that gauge invariance is preserved (recall problem IV.7.3). We will adapt the method on p. 273 for $d = 2 + 1$ spacetime dimensions. For $d = 3 + 1$ dimensions, we had

$$\int d^4p[f(p+a) - f(p)] = \lim_{P \to \infty} ia^\mu \left(\frac{P_\mu}{P}\right) f(P)2\pi^3 P^3$$

where the expression on the right-hand side is to be averaged over the sphere at infinity, for example replacing $P^\mu P^\nu \to \frac{1}{4}\eta^{\mu\nu}P^2$. Adapting this to $d = 2 + 1$ dimensions we have

$$\int d^3p[f(p+a) - f(p)] = \lim_{P \to \infty} ia^\mu \left(\frac{P_\mu}{P}\right) f(P)4\pi P^2 .$$

Replacing $f(p)$ with the integrand

$$f^{\mu\nu}(p) \equiv \text{tr}\left(\gamma^\nu \frac{1}{\not{p}+\not{q}-m}\gamma^\mu \frac{1}{\not{p}-m}\right)$$

we find that the integral $\mathcal{I}^{\mu\nu}(a) \equiv \int \frac{d^3p}{(2\pi)^3} f^{\mu\nu}(p+a)$ satisfies the relation

$$\mathcal{I}^{\mu\nu}(a) = \mathcal{I}^{\mu\nu}(0) + \lim_{P\to\infty} \frac{i}{2\pi^2} a^\mu \left(\frac{P_\mu}{P}\right) f^{\mu\nu}(P) P^2 .$$

In this expression, $f^{\mu\nu}(P)$ simplifies to

$$f^{\mu\nu}(P) = \text{tr}\left(\gamma^\nu \frac{1}{\not{P}}\gamma^\mu \frac{1}{\not{P}}\right) = \frac{\text{tr}(\gamma^\nu \not{P}\gamma^\mu \not{P})}{(P^2)^2} = \frac{2P^\mu P^\nu - \eta^{\mu\nu}P^2}{(P^2)^2}$$

$$= \frac{2(\frac{1}{3}\eta^{\mu\nu}P^2) - \eta^{\mu\nu}P^2}{(P^2)^2} = -\frac{1}{3P^2}\eta^{\mu\nu} .$$

Therefore, our integral satisfies

$$\mathcal{I}^{\mu\nu}(a) = \mathcal{I}^{\mu\nu}(0) - \lim_{P\to\infty} \frac{i}{6\pi^2} a^\lambda \left(\frac{P_\lambda}{P}\right) \eta^{\mu\nu} .$$

The strategy now is to evaluate $\mathcal{I}^{\mu\nu}(0)$ with a regulator in place, then to choose the vector a^μ such that $\mathcal{I}^{\mu\nu}(a)$ is gauge invariant, meaning $q_\mu \mathcal{I}^{\mu\nu}(a) \equiv 0$ and $q_\nu \mathcal{I}^{\mu\nu}(a) \equiv 0$.

Move the gamma matrices into the numerator using $1/(\not{p}-m) = (\not{p}+m)/(p^2-m^2)$ to get

$$\mathcal{I}^{\mu\nu}(0) = \int \frac{d^3p}{(2\pi)^3} \frac{N^{\mu\nu}}{[p^2-m^2][(p+q)^2-m^2]} , \quad N^{\mu\nu} = \text{tr}\left[\gamma^\mu(\not{p}+m)\gamma^\nu(\not{p}+\not{q}+m)\right] .$$

Using the gamma matrix identities and tracing, we find

$$N^{\mu\nu} = 2[p^\mu(p+q)^\nu + p^\nu(p+q)^\mu - \eta^{\mu\nu}p\cdot(p+q) - im\,e^{\mu\nu\lambda}q_\lambda + m^2\eta^{\mu\nu}] .$$

The denominator can be simplified using the Feynman trick

$$\frac{1}{AB} = \int_0^1 dx \frac{1}{[xA + (1-x)B]^2}$$

with $A = p^2 - m^2$ and $B = (p+q)^2 - m^2$. After some arithmetic, this leads to

$$\frac{1}{[p^2-m^2][(p+q)^2-m^2]} = \int_0^1 dx \frac{1}{(\ell^2 + D)^2}$$

where $\ell \equiv p + (1-x)q$ and $D \equiv x(1-x)q^2 - m^2$. The denominator depends on ℓ only as ℓ^2, so we can shift the integration variable from p to ℓ, ignore terms linear in ℓ in the numerator, and make the replacement $\ell^\mu \ell^\nu \to \frac{1}{3}\eta^{\mu\nu}\ell^2$. This leads to

$$N^{\mu\nu} = 2\left[-\frac{1}{3}\eta^{\mu\nu}\ell^2 + x(1-x)(\eta^{\mu\nu}q^2 - 2q^\mu q^\nu) - im\,\varepsilon^{\mu\nu\lambda}q_\lambda + m^2\eta^{\mu\nu}\right] .$$

Now rotate to Euclidean momentum: $\int d^3\ell = i \int d^3\ell_E$ and $\ell^2 = -\ell_E^2$. Also define $D_E \equiv -D = x(1-x)q_E^2 + m^2$ but leave q as Minkowski in the numerator. In total, we have

$$\mathcal{I}^{\mu\nu}(0) = \frac{i}{4\pi^3}\int_0^1 dx \int d^3\ell_E \frac{+\frac{1}{3}\eta^{\mu\nu}\ell_E^2 + K^{\mu\nu}}{(\ell_E^2 + D_E)^2}$$

where

$$K^{\mu\nu} \equiv x(1-x)(\eta^{\mu\nu}q^2 - 2q^\mu q^\nu) - im\,\varepsilon^{\mu\nu\lambda}q_\lambda + m^2\eta^{\mu\nu}$$

does not depend on ℓ_E. The relevant integrals over ℓ_E are

$$\int d^3\ell_E \frac{1}{(\ell_E^2 + D_E)^2} = \frac{\pi^2}{D_E^{1/2}}$$

and

$$\int d^3\ell_E \frac{\ell_E^2}{(\ell_E^2 + D_E)^2} = 2\pi\Lambda - 3\pi^2 D_E^{1/2}$$

where we have regulated the second integral by imposing a cutoff $\Lambda \equiv (\sqrt{q_E^2})_{\max}$. We now have $\mathcal{I}^{\mu\nu}(0) =$

$$\frac{i}{4\pi}\Big\{ \frac{2\Lambda}{3\pi^2}\eta^{\mu\nu} - \eta^{\mu\nu}\int_0^1 dx\, D_E^{+1/2} + (\eta^{\mu\nu}q^2 - 2q^\mu q^\nu)\int_0^1 dx\, x(1-x)D_E^{-1/2}$$

$$+ (-im\,\varepsilon^{\mu\nu\lambda}q_\lambda + m^2\eta^{\mu\nu})\int_0^1 dx\, D_E^{-1/2} \Big\}\,.$$

Now we need the integrals over x. In terms of $\beta \equiv 2m/\sqrt{q_E^2}$, the relevant integrals are:

$$\int_0^1 dx\, D_E^{+1/2} = \frac{m}{2\beta}[(1+\beta^2)\cot^{-1}\beta + \beta]$$

$$\int_0^1 dx\, D_E^{-1/2} = \frac{\beta}{m}\cot^{-1}\beta$$

$$\int_0^1 dx\, x(1-x)\, D_E^{-1/2} = \frac{\beta}{8m}[(1-\beta^2)\cot^{-1}\beta + \beta]$$

We therefore have

$$\mathcal{I}^{\mu\nu}(0) = \frac{i}{4\pi}\left\{ \frac{2\Lambda}{3\pi^2}\eta^{\mu\nu} + \eta^{\mu\nu}R - \frac{\beta}{4m}[(1-\beta^2)\cot^{-1}\beta + \beta]q^\mu q^\nu - i\varepsilon^{\mu\nu\lambda}q_\lambda\,\beta\cot^{-1}\beta \right\}$$

where

$$R \equiv -\frac{m}{2\beta}[(1+\beta^2)\cot^{-1}\beta + \beta] + q^2\frac{\beta}{8m}[(1-\beta^2)\cot^{-1}\beta + \beta] + m\beta\cot^{-1}\beta\,.$$

The term R better simplify in such a way that $q^\mu q^\nu$ and $q^2\eta^{\mu\nu}$ appear only in the combination $q^\mu q^\nu - q^2\eta^{\mu\nu}$ if the integral is to be gauge invariant. Indeed, this will be the case. Since

177

$q_E^2 = 4m^2/\beta^2 = -q^2$, we have $q^2 = -4m^2/\beta^2$ so

$$R = -\frac{m}{2\beta}[(1+\beta^2)\cot^{-1}\beta + \beta] - \left(\frac{4m^2}{\beta^2}\right)\frac{\beta}{8m}[(1-\beta^2)\cot^{-1}\beta + \beta] + m\beta\cot^{-1}\beta$$

$$= -\frac{m}{2\beta}[(1+\beta^2)\cot^{-1}\beta + \beta] - \frac{m}{2\beta}[(1-\beta^2)\cot^{-1}\beta + \beta] + m\beta\cot^{-1}\beta$$

$$= -\frac{m}{2\beta}[(1+\beta^2)\cot^{-1}\beta + \beta + (1-\beta^2)\cot^{-1}\beta + \beta - 2\beta^2\cot^{-1}\beta]$$

$$= -\frac{m}{2\beta}[2\cot^{-1}\beta + 2\beta - 2\beta^2\cot^{-1}\beta]$$

$$= -\frac{m}{\beta}[(1-\beta^2)\cot^{-1}\beta + \beta]$$

Therefore we have

$$\mathcal{I}^{\mu\nu}(0) = \frac{i}{4\pi}\left\{\frac{2\Lambda}{3\pi^2}\eta^{\mu\nu} - \frac{\beta}{4m}[(1-\beta^2)\cot^{-1}\beta + \beta][q^\mu q^\nu + \eta^{\mu\nu}\frac{4m^2}{\beta^2}] - i\varepsilon^{\mu\nu\lambda}q_\lambda\beta\cot^{-1}\beta\right\} .$$

Since $\beta^2 = -4m^2/q^2$, we have

$$\mathcal{I}^{\mu\nu}(0) = \frac{i}{4\pi}\left\{\frac{2\Lambda}{3\pi^2}\eta^{\mu\nu} - \frac{\beta}{4m}[(1-\beta^2)\cot^{-1}\beta + \beta][q^\mu q^\nu - \eta^{\mu\nu}q^2] - i\varepsilon^{\mu\nu\lambda}q_\lambda\beta\cot^{-1}\beta\right\} .$$

The finite part of $\mathcal{I}^{\mu\nu}(0)$ therefore satisfies $q_\mu\mathcal{I}^{\mu\nu}(0) = 0$ and $q_\nu\mathcal{I}^{\mu\nu}(0) = 0$. As discussed previously, we want

$$\mathcal{I}^{\mu\nu}(a) = \mathcal{I}^{\mu\nu}(0) - \lim_{P\to\infty}\frac{i}{6\pi^2}a^\lambda\left(\frac{P_\lambda}{P}\right)\eta^{\mu\nu}$$

to be gauge invariant. We therefore choose a^μ such that

$$\lim_{P\to\infty}a^\mu\left(\frac{P_\mu}{P}\right) = \frac{\Lambda}{\pi}$$

which makes

$$\mathcal{I}^{\mu\nu}(a) = -\frac{i}{4\pi}\left\{\frac{\beta}{4m}[(1-\beta^2)\cot^{-1}\beta + \beta][q^\mu q^\nu - \eta^{\mu\nu}q^2] + i\varepsilon^{\mu\nu\lambda}q_\lambda\beta\cot^{-1}\beta\right\}$$

satisfy $q_\mu\mathcal{I}^{\mu\nu}(a) = 0$ and $q_\nu\mathcal{I}^{\mu\nu}(a) = 0$. Finally, let $\beta = i\alpha$ with $\alpha \equiv 2m/\sqrt{q^2}$ to write the integral in terms of the Minkowskian momentum q^μ. Since

$$\cot^{-1}(i\alpha) = -i\frac{1}{2}\ln\left(\frac{\alpha+1}{\alpha-1}\right)$$

we have

$$\beta[(1-\beta^2)\cot^{-1}\beta + \beta] = \alpha\left[\frac{1}{2}(1+\alpha^2)\ln\left(\frac{\alpha+1}{\alpha-1}\right) - \alpha\right]$$

and

$$\beta\cot^{-1}\beta = \frac{1}{2}\alpha\ln\left(\frac{\alpha+1}{\alpha-1}\right) .$$

Finally, in terms of the variable $\alpha \equiv 2m/\sqrt{q^2}$, the gauge invariant integral is

$$\mathcal{I}^{\mu\nu}(a) = -i\frac{\alpha}{8\pi}\left\{\frac{1}{4m}\left[(1+\alpha^2)\ln\left(\frac{\alpha+1}{\alpha-1}\right) - 2\alpha\right](q^\mu q^\nu - \eta^{\mu\nu}q^2) + i\varepsilon^{\mu\nu\lambda}q_\lambda\ln\left(\frac{\alpha+1}{\alpha-1}\right)\right\} .$$

VI.2 Quantum Hall Fluids

1. To define filling factor precisely, we have to discuss the quantum Hall system on a sphere rather than on a plane. Put a magnetic monopole of strength G (which according to Dirac can be only a half-integer or an integer) at the center of a unit sphere. The flux through the sphere is equal to $N_\phi = 2G$. Show that the single electron energy is given by $E_\ell = (\frac{1}{2}\hbar\omega_c)[\ell(\ell+1)-G^2]/G$ with the Landau levels corresponding to $\ell = G, G+1, G+2, ...,$ and that the degeneracy of the ℓth level is $2\ell + 1$. With L Landau levels filled with non-interacting electrons ($\nu = L$) show that $N_\phi = \nu^{-1}N_e - \mathcal{S}$, where the topological quantity $\mathcal{S}$ is known as the shift.

Solution:

We follow X. G. Wen and A. Zee, "Shift and Spin Vector: New Topological Quantum Numbers for the Hall Fluids," Phys. Rev. Lett., Vol. 69, No. 6, 10 Aug 1992.

The point of this problem is to recognize that since a curved space has a connection 1-form ω (see p. 443), it is possible to write down a Chern-Simons-type interaction between ω and the gauge potential a_μ of the Hall fluid, which we remind the reader is defined as the potential for the conserved electromagnetic current: $J^\mu = \frac{1}{2\pi}\varepsilon^{\mu\nu\lambda}\partial_\nu a_\lambda$. Thus to the Lagrangian of equation (3) on p. 325, we add the interaction term

$$\mathcal{L}_s = s\,\omega_\mu J^\mu = \frac{s}{2\pi}\,\omega_\mu\varepsilon^{\mu\nu\lambda}\partial_\nu a_\lambda$$

where s is a real number. For example, on the sphere the connection 1-form is $\omega^{ab} = -\varepsilon^{ab}\cos\theta\,d\varphi$, where $\varepsilon^{12} = +1$. (See p. 444 in the text.)

Augmenting the discussion on pp. 326-327 with this new term, we find the electromagnetic current

$$J^\mu_{\text{EM}} = -\frac{\partial\mathcal{L}_{\text{eff}}}{\partial A_\mu} = +\frac{1}{2\pi k}\varepsilon^{\mu\nu\lambda}\partial_\nu(A_\lambda - s\omega_\lambda)$$

and the spin current

$$J^\mu_s = \frac{\partial\mathcal{L}_{\text{eff}}}{\partial\omega_\mu} = \frac{s}{2\pi k}\varepsilon^{\mu\nu\lambda}\partial_\nu(-A_\lambda + s\omega_\lambda)\,.$$

The zero component of each current is a number density of each type. Along with the number of electrons $N_e = \int d^2x J^0_{\text{EM}}$ and the number of spin quanta $N_s = \int d^2x J^0_s$, define the number of flux quanta $N_\phi \equiv \frac{1}{2\pi}\int F$ and the number of curvature quanta $N_R \equiv \frac{1}{2\pi}\int R$. Upon recognizing $\varepsilon^{ij}\partial_i A_j d^2x = dA = F$ and $\varepsilon^{ij}\partial_i\omega_j d^2x = d\omega = R$, integrating the $\mu = 0$ components of the currents implies

$$\begin{pmatrix} N_e \\ N_s \end{pmatrix} = \frac{1}{k}\begin{pmatrix} 1 & -s \\ -s & s^2 \end{pmatrix}\begin{pmatrix} N_\phi \\ N_R \end{pmatrix}\,.$$

Inverting this matrix equation shows that

$$N_\phi = kN_e + sN_R\,.$$

In the text it was shown that the filling factor is $\nu = 1/k$ (p. 327), so we have obtained $N_\phi = \nu^{-1}N_e - \mathcal{S}$, with a shift

$$\mathcal{S} = -sN_R = -s\int \frac{R}{2\pi}.$$

The Gauss-Bonnet theorem says that $\frac{1}{2\pi}\int R = 2(1-g)$, where g is the genus of the manifold over which we integrate. In particular, the sphere has genus zero so the shift is $\mathcal{S} = -2s$.

2. For a challenge, derive the effective field theory for Hall fluids with filling factor $\nu = m/k$ with k an odd integer, such as $\nu = \frac{2}{5}$. [Hint: You have to introduce m gauge potentials $a_{I\lambda}$ and generalize (2) to $J^\mu = (1/2\pi)\varepsilon^{\mu\nu\lambda}\partial_\nu \sum_{I=1}^m a_{I\lambda}$. The effective theory turns out to be

$$\mathcal{L} = \frac{1}{4\pi}\sum_{I,J=1}^m a_I K_{IJ}\varepsilon\partial a_J + \sum_{I=1}^m a_{I\mu}\tilde{j}^{I\mu} + ...$$

with the integer k replaced by a matrix K. Compare with (13).]

$$(13) \qquad \mathcal{L} = \sum_{I,J}\frac{1}{4\pi}K_{IJ}a_I\varepsilon\partial a_J + ...$$

Solution:

For this problem and the next, we follow J. Fröhlich and A. Zee, "Large Scale Physics of the Quantum Hall Fluid," Nucl. Phys. B364 (1991) 517-540.

For notational convenience we sometimes use differential forms and work with the Lagrangian volume-form $L \equiv 4\pi\mathcal{L}\,d^3x$. The gauge potential a_μ for the conserved electromagnetic current J^μ is defined by the equation $J = \frac{1}{2\pi}{}^*da$, where * denotes the Hodge-star operator that takes p-forms to $(3-p)$-forms (see IV.4.2). The Chern-Simons Lagrangian $\mathcal{L} = \frac{1}{4\pi}\varepsilon^{\mu\nu\lambda}a_\mu\partial_\nu a_\lambda$ is written $L = a\,da$. (We set the coefficient of the Chern-Simons theory that results from the long-distance physics of the non-interacting Hall fluid to 1.)

Consider a 2D system of non-interacting electrons with m Landau levels filled: $\nu = \frac{2\pi n_e}{B} = m$. The large gap between Landau levels implies that it is reasonable to take the levels as dynamically independent, so that the current of electrons belonging to each Landau level is separately conserved.

Let $I = 1, 2, ..., m$ label the m levels. For each conserved electromagnetic current J_I, we introduce a gauge potential a_I defined by $J_I = \frac{1}{2\pi}{}^*da_I$. The new Lagrangian is $L = \sum_{I=1}^m a_I da_I$.

Now turn on interactions between the electrons, which couple the different Landau levels. As discussed in the text, the resulting long-distance theory will be described by another

180

Chern-Simons Lagrangian. The interactions should only involve the total electromagnetic current

$$J = \sum_{I=1}^{m} J_I = \frac{1}{2\pi} * \left(\sum_{I=1}^{m} da_I \right)$$

not the individual currents J_I. As a result, electron-electron interactions can only change the Lagrangian by adding a Cern-Simons term of the form $(\sum_I a_I)(\sum_J da_J)$. Thus, the new Lagrangian is

$$L = \sum_{I=1}^{m} a_I da_I + p \sum_{I,J=1}^{m} a_I da_J \equiv \sum_{I,J=1}^{m} a_I K_{IJ} da_J$$

with p a real number. Here we have defined the m-by-m matrix

$$K = \begin{pmatrix} p+1 & p & \cdots & p \\ p & p+1 & & \vdots \\ \vdots & & \ddots & p \\ p & \cdots & p & p+1 \end{pmatrix}.$$

Unpackaging the forms notation, we have arrived at the Lagrangian

$$\mathcal{L}_0 = \frac{1}{4\pi} \sum_{I,J=1}^{m} \varepsilon^{\mu\nu\lambda} a_{I\mu} K_{IJ} \partial_\nu a_{J\lambda} .$$

Denote the current of quasiparticles (the "vortex current") in the Ith Landau level by j_I^μ. This couples to the Chern-Simons gauge potential of each level as

$$\mathcal{L}_1 = \sum_{I=1}^{m} a_{I\mu} j_I^\mu$$

by definition. ("Here we define the quasiparticles as the entities that couple to the gauge potential..." above equation (5) on p. 326.)

Now subject the system to additional external electromagnetic fields described by the vector potential $\mathcal{A}_\mu$. This couples to the Chern-Simons gauge potentials through the electromagnetic current interaction for each level:

$$\mathcal{L}_2 = \sum_{I=1}^{m} J_I^\mu \mathcal{A}_\mu = \sum_{I=1}^{m} \left(\frac{1}{2\pi} \varepsilon^{\mu\nu\lambda} \partial_\nu a_{I\lambda} \right) \mathcal{A}_\mu = -\sum_{I=1}^{m} \frac{1}{2\pi} \varepsilon^{\mu\nu\lambda} a_{I\lambda} \partial_\nu \mathcal{A}_\mu$$

where we have dropped a total derivative. The full Lagrangian $\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_1 + \mathcal{L}_2$ is

$$\mathcal{L} = \frac{1}{4\pi} \sum_{I,J=1}^{m} \varepsilon^{\mu\nu\lambda} a_{I\mu} K_{IJ} \partial_\nu a_{J\lambda} + \sum_{I=1}^{m} a_{I\mu} \left(j_I^\mu - \frac{1}{2\pi} \varepsilon^{\mu\nu\lambda} \partial_\nu \mathcal{A}_\lambda \right) .$$

181

The effective current $\tilde{j}_I^\mu = j_I^\mu - \frac{1}{2\pi}\varepsilon^{\mu\nu\lambda}\partial_\nu\mathcal{A}_\lambda$ is called the "reduced vorticity current." To understand the long-distance physics described by $\mathcal{L}$, we integrate out the gauge fields a_I to obtain a matrix-valued version of the Hopf Lagrangian of equation (6) on p. 326:

$$\mathcal{L}_{\text{eff}} = \pi \sum_{I,J=1}^{m} \tilde{j}_{I\mu}(K^{-1})_{IJ}\varepsilon^{\mu\nu\lambda}\partial_\nu\left(\frac{1}{\partial^2}\right)\tilde{j}_{J\lambda} \,.$$

The inverse of the matrix K is

$$K^{-1} = \begin{pmatrix} 1 - \frac{p}{1+mp} & -\frac{p}{1+mp} & \cdots & -\frac{p}{1+mp} \\ -\frac{p}{1+mp} & 1 - \frac{p}{1+mp} & & \\ \vdots & & \ddots & \\ -\frac{p}{1+mp} & & & 1 - \frac{p}{1+mp} \end{pmatrix}$$

with $1 - p/(1+mp)$ on the diagonals and $-p/(1+mp)$ everywhere else. This is the field theory whose consequences we will study in the next problem.

3. For the Lagrangian in (13), derive the analogs of (8), (9), and (11).

$$(8) \qquad J_{\text{em}}^\mu = \frac{1}{4\pi k}\varepsilon^{\mu\nu\lambda}\partial_\nu A_\lambda$$

$$(9) \qquad \mathcal{L} = \frac{1}{k}A_\mu j^\mu$$

$$(11) \qquad \frac{\theta}{\pi} = \frac{1}{k}$$

Solution:

In the previous problem, VI.2.2, we derived the effective Lagrangian

$$\mathcal{L}_{\text{eff}} = \pi \sum_{I,J=1}^{m} \tilde{j}_{I\mu}(K^{-1})_{IJ}\varepsilon^{\mu\nu\lambda}\partial_\nu\left(\frac{1}{\partial^2}\right)\tilde{j}_{J\lambda}$$

with the matrix

$$K^{-1} = \begin{pmatrix} 1 - \frac{p}{1+mp} & -\frac{p}{1+mp} & \cdots & -\frac{p}{1+mp} \\ -\frac{p}{1+mp} & 1 - \frac{p}{1+mp} & & \\ \vdots & & \ddots & \\ -\frac{p}{1+mp} & & & 1 - \frac{p}{1+mp} \end{pmatrix} \,.$$

This Lagrangian $\mathcal{L}_{\text{eff}}$ contains three types of terms: $\mathcal{A}\mathcal{A}$, $\mathcal{A}j$ and jj. The self-interaction of the electromagnetic field is given by

$$\mathcal{L}_{\text{eff}}^{(\mathcal{A}\mathcal{A})} = -\frac{1}{4\pi}\left(\sum_{I,J=1}^{m}(K^{-1})_{IJ}\right)\varepsilon^{\mu\nu\lambda}\mathcal{A}_\mu\partial_\nu\mathcal{A}_\lambda \,.$$

182

Varying this with respect to $\mathcal{A}$ gives the expectation value of the total electromagnetic current:

$$\langle J^\mu_{\text{EM}} \rangle = \frac{1}{2\pi} \left(\sum_{I,J=1}^{m} (K^{-1})_{IJ} \right) \varepsilon^{\mu\nu\lambda} \partial_\nu \mathcal{A}_\lambda \ .$$

As explained on p. 327 of the text, the $\mu = 0$ component of this equation implies that the filling factor is

$$\nu = \sum_{I,J=1}^{m} (K^{-1})_{IJ} = \frac{m}{1+mp} = \frac{1}{1+p}, \ \frac{2}{1+2p}, \ \frac{3}{1+3p}, \ \dots$$

This is what replaces k^{-1} in equation (8). Thus for $p = 2$ we have derived the theory of quantum Hall fluids with filling factors

$$\nu = \frac{1}{3}, \ \frac{2}{5}, \ \frac{3}{7}, \ \dots$$

which realizes the claim of problem VI.2.2.

From the spatial components of $\langle J^\mu_{\text{EM}} \rangle \propto \varepsilon^{\mu\nu\lambda} \partial_\nu \mathcal{A}_\lambda$, we also learn that an electric field in the y-direction produces a current in the x-direction with a proportionality constant $\sigma_H = \sum_{I,J=1}^{m} (K^{-1})_{IJ} = \nu$, known as the Hall conductance.

Next we want to compute the electric charge of each quasiparticle. To do this, consider the coupling of the quasiparticle currents j_I to the applied gauge field $\mathcal{A}$:

$$\mathcal{L}^{\mathcal{A}j}_{\text{eff}} = \mathcal{A}_\mu \sum_{I,J=1}^{m} (K^{-1})_{IJ} j^\mu_J \equiv \mathcal{A}_\mu J^\mu_{\text{EM}}$$

where we have identified the total electromagnetic current induced by a quasiparticle currents j^μ_I. (All of these currents are actually current densities, but as is common we are being sloppy with the language.)

Recall from earlier that the total electromagnetic current is comprised of a sum of m individually conserved electromagnetic currents for each Landau level:

$$J^\mu_{\text{EM}} = \sum_{I=1}^{m} J^\mu_I = \sum_{I=1}^{m} \frac{1}{2\pi} \varepsilon^{\mu\nu\lambda} \partial_\nu a_{I\lambda} \ .$$

Varying the total Lagrangian

$$\mathcal{L} = \frac{1}{4\pi} \sum_{I,J=1}^{m} \varepsilon^{\mu\nu\lambda} a_{I\mu} K_{IJ} \partial_\nu a_{J\lambda} + \sum_{I=1}^{m} a_{I\mu} \left(j^\mu_I - \frac{1}{2\pi} \varepsilon^{\mu\nu\lambda} \partial_\nu \mathcal{A}_\lambda \right)$$

with respect to $a_{I\mu}$ and recognizing $\frac{1}{2\pi} \varepsilon^{\mu\nu\lambda} \partial_\nu a_{I\lambda} = J^\mu_I$, we obtain the matrix equation

$$j^\mu_I = \sum_{J=1}^{m} K_{IJ} J^\mu_J \ .$$

183

Since we have already computed the matrix inverse of K, this can be inverted immediately to obtain

$$J_I^\mu = \sum_{J=1}^m (K^{-1})_{IJ} j_J^\mu \ .$$

The $\mu = 0$ component of this equation determines the charge density, which we can integrate to obtain the total electric charge induced in the Ith Landau level:

$$q_I \equiv \int d^2x\, J_I^0 = \sum_{J=1}^m (K^{-1})_{IJ} \Phi_J$$

where we have defined the "vorticities" $\Phi_I = \int d^2x\, j_I^0$. In this language, q_I is the electric charge bound to the vorticity Φ_I in the Ith Landau band. Using the explicit form of K^{-1}, the charge is

$$q_I = \left(1 - \frac{p}{1+mp}\right)\Phi_I - \frac{p}{1+mp} \sum_{J=1,\ J\neq I}^m \Phi_J$$

$$= \Phi_I - \frac{p}{1+mp} \sum_{J=1}^m \Phi_J \ .$$

For $\Phi_I = 1$, $I = 1, ..., m$ we find

$$q_I = \frac{1}{1+mp} \ .$$

This is the replacement of $q = 1/k$ from equation (9).

Now let us turn to the jj interaction term:

$$\mathcal{L}_{\text{eff}}^{(jj)} = \pi \sum_{I,J=1}^m \varepsilon_{\mu\nu\lambda} j_I^\mu (K^{-1})_{IJ} \partial^\nu \left(\frac{1}{\partial^2}\right) j_J^\lambda \ .$$

As in the case of just one filled Landau level, this interaction describes a generalized version of the Aharonov-Bohm effect. An excitation of the system with vorticity vector $(\Phi_1, ..., \Phi_m)$ results in a statistics phase

$$\frac{\theta}{\pi} = \sum_{I,J=1}^m \Phi_I (K^{-1})_{IJ} \Phi_J = \frac{m}{1+mp}$$

where in the last line we have used $\Phi_I = 1$ for $I = 1, ..., m$. This is the generalization of equation (11).

184

VI.4 The σ Models as Effective Field Theories

1. Show that the vacuum expectation value of $(\sigma, \vec{\pi})$ can indeed point in any direction without changing the physics. At first sight, this statement seems strange since, by virtue of its γ_5 coupling to the nucleon, the pion is a pseudoscalar field and cannot have a vacuum expectation value without breaking parity. But $(\sigma, \vec{\pi})$ are just Greek letters. Show that by a suitable transformation of the nucleon field parity is conserved, as it should be in the strong interaction.

Solution:

Before proceeding to the solution, it is useful to review the sigma model for nucleons and pions using slightly different notation from the text.

The nucleon Lagrangian is

$$\mathcal{L} = g\,\bar{\psi}(\sigma I + i\vec{\pi}\cdot\vec{\tau}\gamma_5)\psi \;=\; g\,\bar{\psi}_L(\sigma I + i\vec{\pi}\cdot\vec{\tau})\psi_R + h.c.$$

This problem is about transformation properties under the global $SU(2)_L \otimes SU(2)_R$ symmetry whose diagonal piece $SU(2)_I \subset SU(2)_L \otimes SU(2)_R$ is Heisenberg's isospin. To remove possible sources of confusion about the Dirac spinor indices and to unclutter the notation as little as possible, let us write the Dirac field as two-component spinors: $\psi = (\chi, \bar{\chi}^\dagger)^T$. (Review Appendix E now if you are unfamiliar with this notation.) The above Lagrangian can be rewritten as

$$\mathcal{L} = -g\,\chi_a \varepsilon^{ab}\Pi_{b\dot{c}}\bar{\chi}^{\dot{c}} + h.c.$$

where we have defined the matrix

$$\Pi_{b\dot{c}} \equiv \sum_{\mu=1}^{4}\frac{1}{2}\pi_\mu(\tau^\mu)_{b\dot{c}} = \frac{1}{2}\begin{pmatrix} \sigma + i\pi_3 & i(\pi_1 - i\pi_2) \\ i(\pi_1 + i\pi_2) & \sigma - i\pi_3 \end{pmatrix}_{b\dot{c}}.$$

Here $\{\tau^\mu\}_{\mu=1}^3$ are the Pauli matrices times a factor of i:

$$\tau^1 \equiv i\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau^2 \equiv i\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau^3 \equiv i\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Also, $\tau^4 \equiv I$ is the 2×2 identity matrix, and $\pi_4 \equiv \sigma$ is the fourth meson field. (Review Appendix B for the local isomorphism $SU(2) \otimes SU(2) \simeq SO(4)$, which necessitates the inclusion of the fourth meson field.)

The factors of $\frac{1}{2}$ give the correctly normalized kinetic term $\mathcal{L} = \text{tr}(\Pi\Pi^\dagger) = \frac{1}{2}[(\partial\sigma)^2 + (\partial\pi_3)^2] + \partial\pi^+\partial\pi^-$, where $\pi^\pm \equiv \frac{1}{\sqrt{2}}(\pi_1 \mp i\pi_2)$.

Under $SU(2)_L$, we have $\chi_a \to (L\chi)_a = L_a{}^b\chi_b = \chi_b L_a{}^b = \chi_b(L^T)^b{}_a$ and $\Pi_{b\dot{c}} \to L_b{}^c\Pi_{c\dot{c}}$, where L is a 2×2 unitary matrix with determinant 1. Since $\vec{\sigma}^T\varepsilon = -\varepsilon\vec{\sigma}$, we have $L^T\varepsilon = -\varepsilon L^\dagger$. So $SU(2)_L$ acts as $\chi\varepsilon\Pi \to \chi\varepsilon L^\dagger L\Pi = \chi\varepsilon\Pi$, or in other words the Lagrangian is

invariant under $SU(2)_L$. Under $SU(2)_R$, we have $\bar{\chi}^{\dot{c}} \to R^{\dot{c}}{}_{\dot{e}}\bar{\chi}^{\dot{e}}$ and $\Pi_{b\dot{c}} \to \Pi_{b\dot{e}}(R^{-1})^{\dot{e}}{}_{\dot{c}}$, so $\Pi\bar{\chi} \to \Pi R^{-1}R\bar{\chi} = \Pi\bar{\chi}$, so the Lagrangian is invariant under $SU(2)_R$. This is what is meant by stating that the Lagrangian is invariant under $SU(2)_L \otimes SU(2)_R$.

Now it is clear that the vacuum alignment may be chosen to point in an arbitrary direction in $SO(4) \simeq SU(2)_L \otimes SU(2)_R$, since we can always perform $SU(2)_L \otimes SU(2)_R$ transformations on the nucleon fields to leave the physics invariant. In other words, the vacuum expectation values $\langle \pi_\mu \rangle$ just need to satisfy the constraint

$$\sum_{\mu=0}^{3} \langle \pi_\mu \rangle^2 = \langle \sigma \rangle^2 + \sum_{i=1}^{3} \langle \pi_i \rangle^2 \equiv v^2$$

with v some constant with dimensions of mass. The text chooses $\langle \sigma \rangle = v$ and $\langle \pi_i \rangle = 0$, so that upon spontaneous symmetry breaking the Lagrangian contains the term $\mathcal{L} = gv\,\chi\bar{\chi} + h.c. = gv\,\bar{\psi}_L\psi_R + h.c. = gv\,\bar{\psi}\psi$, which implies a mass $M = gv$ for the nucleons. But any other choice of vacuum alignment can be transformed into this form using the $SU(2)_L \otimes SU(2)_R$ transformation $\langle \Pi \rangle \to L\langle \Pi \rangle R^{-1}$, as long as we also redefine the nucleon fields by $\chi \to L\chi$ and $\bar{\chi} \to R\bar{\chi}$.

For example, choose $\langle \pi_2 \rangle = v$ and $\langle \sigma \rangle = \langle \pi_1 \rangle = \langle \pi_3 \rangle = 0$. Then

$$\langle \Pi \rangle = \begin{pmatrix} 0 & v \\ -v & 0 \end{pmatrix}.$$

Now choose the matrices

$$L = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = -\tau_x \text{ and } R = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \tau_z.$$

The transformation $\langle \Pi \rangle \to L\langle \Pi \rangle R^\dagger$ yields

$$\begin{pmatrix} 0 & v \\ -v & 0 \end{pmatrix} \to \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & v \\ -v & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} v & 0 \\ 0 & v \end{pmatrix}$$

So upon the field redefinitions $\chi \to -\tau_x\chi$ and $\bar{\chi} \to \tau_z\bar{\chi}$, we recover the mass term $\mathcal{L} = gv\,\chi\bar{\chi} + h.c. = gv\,\bar{\psi}_L\psi_R + h.c. = gv\,\bar{\psi}\psi$, as we must by $SO(4) \simeq SU(2)_L \otimes SU(2)_R$ invariance.

VI.5   Ferromagnets and Antiferromagnets

1. Work out the two branches of the spin wave spectrum in the ferromagnetic case, paying particular attention to the polarization.

Solution:

For this problem and the next we follow X. G. Wen and A. Zee, "Spin Waves and Topological Terms in the Mean-Field Theory of Two-Dimensional Ferromagnets and Antiferromagnets," Phys. Rev. Lett., Vol. 61 No. 8, 22 Aug 1988.

The equations of motion are[14]

$$\begin{pmatrix} -\frac{\omega^2}{g^2} + h(\vec{k}) & -\frac{1}{2}i\omega \\ +\frac{1}{2}i\omega & -\frac{\omega^2}{g^2} + h(\vec{k}) \end{pmatrix} \begin{pmatrix} \xi^x(\vec{k},\omega) \\ \xi^y(\vec{k},\omega) \end{pmatrix} = 0 \ .$$

where $h(\vec{k}) \approx 2|J|a^2k^2$ for small $k \equiv |\vec{k}|$ (and $J < 0$). Setting the determinant of the matrix to zero gives the quadratic equation

$$\omega^2 \pm \frac{g^2}{2}\omega - g^2 h(\vec{k}) = 0 \implies \omega^{\mp}(k) = \mp\frac{g^2}{4} + \frac{g^2}{4}\sqrt{1 + \frac{16}{g^2}h(k)} \ .$$

We have chosen the $+$ root in the quadratic equation to keep only positive-frequency solutions. For small k, we have:

$$\omega^+ = +\frac{g^2}{4} + \frac{g^2}{4}\sqrt{1 + \frac{16}{g^2}h(k)} = +\frac{g^2}{2} + O(k^2)$$

$$\omega^- = -\frac{g^2}{4} + \frac{g^2}{4}\sqrt{1 + \frac{16}{g^2}h(k)} \approx +2h(k) = +4|J|a^2k^2$$

so that there is a high frequency branch $\omega^+ = g^2/2 + O(k^2)$ and a low-frequency branch $\omega^- \propto k^2$, which is the typical dispersion relation for a nonrelativistic particle.

Now let us work out the polarizations for each branch. Plugging in $\omega^+ \approx \frac{1}{2}g^2$, dropping all terms of order k^2 and dividing through by $-g^2/4$, the equations of motion become

$$\begin{pmatrix} 1 & +i \\ -i & 1 \end{pmatrix} \begin{pmatrix} \xi^x \\ \xi^y \end{pmatrix} = 0 \ .$$

This is satisfied by $(\xi^x, \xi^y) \propto (1, +i)$. Plugging in $\omega^- \approx 4|J|a^2k^2$ and $h(k) \approx 2|J|a^2k^2$, dropping terms of higher order in k, and dividing through by $2|J|a^2k^2$, the equations of motion become

$$\begin{pmatrix} 1 & -i \\ +i & 1 \end{pmatrix} \begin{pmatrix} \xi^x \\ \xi^y \end{pmatrix} = 0 \ .$$

This is satisfied by $(\xi^x, \xi^y) \propto (1, -i)$.

---

[14]The text uses the notation $\delta n_i \equiv \xi_i$. We choose to change notation in order to distinguish linearizing about the ground state from the arbitrary variation $\vec{n}_i \rightarrow \vec{n}_i + \delta\vec{n}_i$ used to obtain the equations of motion. But of course you are free to use whichever notation you like.

2. Verify that in the antiferromagnetic case the Berry's phase term merely changes the spin wave velocity and does not affect the spectrum qualitatively as in the ferromagnetic case.

As given on p. 346, linearizing around the ground state $\vec{n}_i(t) = (-1)^i \hat{e}_z + \vec{\xi}_i(t)$ implies[15]

$$\begin{pmatrix} -\frac{\omega^2}{g^2} + f(\vec{k}) & -\frac{1}{2}i\omega \\ +\frac{1}{2}i\omega & -\frac{\omega^2}{g^2} + f(\vec{k} + \vec{Q}) \end{pmatrix} \begin{pmatrix} \xi^x(\vec{k}, \omega) \\ \xi^y(\vec{k} + \vec{Q}, \omega) \end{pmatrix} = 0$$

where $f(\vec{k}) = 4J[2 + \sum_{\mu=1}^2 \cos(k_\mu a)]$ and $J > 0$. Note that since $Q_\mu = \frac{\pi}{a}$, and $\cos(x + \pi) = -\cos x$, we have $f(\vec{k} + \vec{Q}) = 4J[2 + \sum_\mu \cos(k_\mu a + \pi)] = 4J[2 - \sum_\mu \cos(k_\mu a)]$. For small k_μ, we have

$$f(\vec{k}) \approx 16J - 2Ja^2k^2 \text{ and } f(\vec{k} + \vec{Q}) \approx +2Ja^2k^2$$

where $k \equiv |\vec{k}|$.

Returning to the equation of motion, setting the determinant of the matrix equal to zero results in a quadratic equation for ω^2:

$$[\omega^2 - g^2 f(\vec{k})][\omega^2 - g^2 f(\vec{k} + \vec{Q})] - \frac{1}{4}g^4\omega^2 = 0 .$$

This has the solutions:

$$\omega_\pm^2 = \frac{g^2}{2}\left[f(\vec{k}) + f(\vec{k} + \vec{Q}) + \frac{1}{4}g^2\right]\left\{1 \pm \sqrt{1 - \frac{4f(\vec{k})f(\vec{k} + \vec{Q})}{\left[f(\vec{k}) + f(\vec{k} + \vec{Q}) + \frac{1}{4}g^2\right]^2}}\right\} .$$

For small k, we find $f(\vec{k}) + f(\vec{k} + \vec{Q}) = 16J + O(k^4)$ and $f(\vec{k})f(\vec{k} + \vec{Q}) = 32J^2a^2k^2 + O(k^4)$. The solutions simplify to:

$$\omega_\pm^2 \approx \frac{g^2}{2}\left[16J + \frac{1}{4}g^2\right]\left\{1 \pm \left(1 - \frac{64J^2a^2k^2}{[16J + \frac{1}{4}g^2]^2}\right)\right\} .$$

The spectrum splits into two branches:

$$\omega_+^2 \approx g^2(16J + \frac{1}{4}g^2) + O(k^2)$$

$$\omega_-^2 \approx \left(\frac{32J^2a^2g^2}{16J + \frac{1}{4}g^2}\right) k^2 .$$

The low-frequency branch is a linear dispersion relation $\omega \propto k$ with spin-wave velocity

$$c = \frac{32J^2a^2g^2}{16J + \frac{1}{4}g^2} .$$

---

[15]There is a typo in equation (6) in the text. See the addendum at the end of this section.

The question now is to track the effect of the Berry's phase term. Without this term (namely the $\pm i\omega$ terms in the equations of motion), setting the determinant of the matrix to zero in the equations of motion implies

$$[\omega^2 - g^2 f(\vec{k})][\omega^2 - g^2 f(\vec{k} + \vec{Q})] = 0$$

which is still a quadratic equation for ω^2. The solutions are $\omega^2 = g^2 f(\vec{k}) \approx g^2(16J - 2Ja^2k^2)$ and $\omega^2 = g^2 f(\vec{k} + \vec{Q}) \approx g^2(2Ja^2k^2)$. Again, there is a high-frequency branch

$$\omega^2_{\text{high}} = 16g^2 J + O(k^2)$$

and a low-frequency branch

$$\omega^2_{\text{low}} = 2Ja^2g^2k^2$$

which is a linear dispersion relation with spin-wave velocity

$$c = 2Ja^2g^2 .$$

We see that the spectrum is qualitatively the same, and the Berry's phase only changes the particular values of the parameters in the dispersion relations.

Addendum: Deriving the Equations of Motion

Here we fill in the steps absent from the text in order to derive the equations of motion. The action is $S = \int dt\, L$ with Lagrangian

$$L = \sum_{i=1}^{\mathcal{N}} \left(i\, z_i^\dagger \partial_t z_i + \frac{1}{2g^2} \partial_t \vec{n}_i \cdot \partial_t \vec{n}_i \right) - J \sum_{\langle ij \rangle} \vec{n}_i \cdot \vec{n}_j$$

where $\vec{n}_i = z_i^\dagger \vec{\sigma} z_i$ is the Pauli-Hopf map introduced in the text, and the sum over nearest neighbors can be written explicitly as

$$\sum_{\langle ij \rangle} = \frac{1}{2} \sum_{i=1}^{\mathcal{N}} \sum_{j=1}^{\mathcal{N}} \sum_{\mu=1}^{d} \left(\delta_{i,\, j+\hat{e}_\mu} + \delta_{i,\, j-\hat{e}_\mu} \right) .$$

The number of lattice sites is $\mathcal{N}$, and the number of lattice spatial dimensions is d. (We specialize to $d = 2$.) The vector $\hat{e}_\mu$ is a unit vector pointing in the μth direction, and the kronecker deltas are to be interpreted according to the example

$$\sum_{i=1}^{\mathcal{N}} \delta_{i,\, j+\hat{e}_\mu} f(\vec{x}_i) = f(\vec{x}_{j+\hat{e}_\mu}) \equiv f(\vec{x}_j + a\, \hat{e}_\mu)$$

where a is the lattice spacing. Using the variation

$$\int dt\, \delta \left(z_i^\dagger \partial_t z_i \right) = \int dt\, \frac{i}{2} \delta \vec{n}_i \cdot (\vec{n}_i \times \partial_t \vec{n}_i)$$

given in the text, we have

$$\delta S = \int dt \sum_{i=1}^{\mathcal{N}} \delta \vec{n}_i \cdot \left[-\frac{1}{2} \vec{n}_i \times \partial_t \vec{n}_i - \frac{1}{g^2} \partial_t^2 \vec{n}_i - J \sum_{j=1}^{\mathcal{N}} \sum_{\mu=1}^{d} \left(\delta_{i,\,j+\hat{e}_\mu} + \delta_{i,\,j-\hat{e}_\mu} \right) \vec{n}_j \right] .$$

When deriving local equations of motion by setting $\delta S = 0$, we must assume that the variations $\delta \vec{n}_i(t)$ are arbitrary functions of i and t – that is, that the sum over i and the integral over t are unconstrained. But here the sum is constrained by $\vec{n}_i \cdot \vec{n}_i = 1$. We need to incorporate this constraint into the path integral:

$$\mathcal{Z} = \int \mathcal{D}n \prod_{i=1}^{\mathcal{N}} \delta(\vec{n}_i \cdot \vec{n}_i - 1)\, e^{i \int dt\, L}$$

$$= \int \mathcal{D}n \prod_{i=1}^{\mathcal{N}} \int \mathcal{D}\nu\, e^{i \int dt\, (\vec{n}_i \cdot \vec{n}_i - 1)\nu}\, e^{i \int dt\, L}$$

$$= \int \mathcal{D}n \mathcal{D}\nu\, e^{i \int dt\, \left[\sum_{i=1}^{\mathcal{N}} (\vec{n}_i \cdot \vec{n}_i - 1)\nu + L \right]}$$

So the total unconstrained Lagrangian is

$$L_{\text{tot}} = L + \sum_{i=1}^{\mathcal{N}} \vec{n}_i \cdot \vec{n}_i \nu - \mathcal{N}\nu .$$

To the previous variation we therefore add the term $2 \sum_{i=1}^{\mathcal{N}} \delta \vec{n}_i \cdot \vec{n}_i\, \nu$, so that the full varied action is

$$\delta S_{\text{tot}} = \int dt \sum_{i=1}^{\mathcal{N}} \delta \vec{n}_i \cdot \left\{ -\frac{1}{2} \vec{n}_i \times \partial_t \vec{n}_i - \frac{1}{g^2} \partial_t^2 \vec{n}_i - \sum_{j=1}^{\mathcal{N}} \left[J \sum_{\mu=1}^{d} \left(\delta_{i,\,j+\hat{e}_\mu} + \delta_{i,\,j-\hat{e}_\mu} \right) - 2\nu \delta_{ij} \right] \vec{n}_j \right\}$$

Now the sum over i is totally unconstrained, so the equation of motion is the term in $\{...\}$ set equal to zero:

$$\frac{1}{2} \vec{n}_i \times \partial_t \vec{n}_i + \frac{1}{g^2} \partial_t^2 \vec{n}_i + \sum_{j=1}^{\mathcal{N}} \mathcal{J}_{ij} \vec{n}_j = 0$$

with

$$\mathcal{J}_{ij} \equiv J \sum_{\mu=1}^{d} \left(\delta_{i,\,j+\hat{e}_\mu} + \delta_{i,\,j-\hat{e}_\mu} \right) - 2\nu \delta_{ij} .$$

First consider the ferromagnet $(J < 0)$. The ground state of the ferromagnet is the state for which all spins point in the positive z-direction. Linearize about this state as $\vec{n}_i(t) = \hat{e}_z + \vec{\xi}_i(t)$ to obtain

$$\vec{n}_i \times \partial_t \vec{n}_i = (\hat{e}_z + \vec{\xi}_i) \times \partial_t(\hat{e}_z + \vec{\xi}_i) = \hat{e}_z \times \partial_t \vec{\xi}_i + O(\xi^2) .$$

As mentioned in the text, the constraint $\vec{n}_i^2 = (\hat{e}_z + \vec{\xi}_i)^2 = 1 + 2\,\hat{e}_z \cdot \vec{\xi}_i + O(\xi^2) = 1$ implies $\hat{e}_z \cdot \vec{\xi}_i = 0$, so that the linearized deviation from the ground state occurs only in the (xy)-plane.

Since $\vec{\xi}$ lies purely in the (x, y)-plane, the equation of motion in the z-direction is:

$$\sum_{j=1}^{\mathcal{N}} \mathcal{J}_{ij} = 0 \implies \nu = Jd \ .$$

This fixes the value of the Lagrange multiplier ν.

In the (x, y)-plane, the equations of motion are:

$$\frac{1}{2}\hat{e}_z \times \partial_t \vec{\xi}_i(t) + \frac{1}{g^2}\partial_t^2 \vec{\xi}_i(t) + \sum_{j=1}^{\mathcal{N}} \mathcal{J}_{ij}\vec{\xi}_j(t) = 0$$

where now

$$\mathcal{J}_{ij} = J\left[\sum_{\mu=1}^{d}\left(\delta_{i,j+\hat{e}_\mu} + \delta_{i,j-\hat{e}_\mu}\right) - 2d\,\delta_{ij}\right]$$

and as usual $d = 2$.

Now that the equation of motion is linear in $\vec{\xi}_i(t)$, we may Fourier transform using the convention

$$f(\vec{k}, \omega) \equiv \sum_{i=1}^{\mathcal{N}} \int_{-\infty}^{\infty} dt\, e^{-i(\omega t + \vec{k}\cdot\vec{x}_i)} f_i(t) \ .$$

We have

$$\sum_{i=1}^{\mathcal{N}} \int_{-\infty}^{\infty} dt\, e^{-i(\omega t + \vec{k}\cdot\vec{x}_i)}\partial_t^2 \vec{\xi}_i(t) = -\omega^2\,\vec{\xi}(\vec{k}, \omega)$$

and

$$\sum_{i=1}^{\mathcal{N}} \int_{-\infty}^{\infty} dt\, e^{-i(\omega t + \vec{k}\cdot\vec{x}_i)}\sum_{j=1}^{\mathcal{N}} \mathcal{J}_{ij}\vec{\xi}_j = \sum_{i=1}^{\mathcal{N}} \int_{-\infty}^{\infty} dt\, e^{-i(\omega t + \vec{k}\cdot\vec{x}_i)}J\left[\sum_{\mu=1}^{d}\left(\delta_{i,j+\hat{e}_\mu} + \delta_{i,j-\hat{e}_\mu}\right) - 2d\,\delta_{ij}\right]\vec{\xi}_j(t)$$

$$= \sum_{j=1}^{\mathcal{N}} \int_{-\infty}^{\infty} dt\, e^{-i\omega t}J\left[\sum_{\mu=1}^{d}\left(e^{-i\vec{k}\cdot(\vec{x}_j + a\hat{e}_\mu)} + e^{-i\vec{k}\cdot(\vec{x}_j - a\hat{e}_\mu)}\right) - 2d\,e^{-i\vec{k}\cdot\vec{x}_j}\right]\vec{\xi}_j(t)$$

$$= J\left[\sum_{\mu=1}^{d}\left(e^{-i\vec{k}\cdot\hat{e}_\mu a} + e^{+i\vec{k}\cdot\hat{e}_\mu a}\right) - 2d\right]\sum_{j=1}^{\mathcal{N}} \int_{-\infty}^{\infty} dt\, e^{-i(\omega t + \vec{k}\cdot\vec{x}_j)}\vec{\xi}_j(t)$$

$$= 2J\left[\sum_{\mu=1}^{d}\cos(k_\mu a) - d\right]\vec{\xi}(\vec{k}, \omega) \quad , \quad k_\mu \equiv \vec{k}\cdot\hat{e}_\mu \ .$$

For the Berry's phase term, we have:

$$\sum_{i=1}^{\mathcal{N}} \int_{-\infty}^{\infty} dt\, e^{-i(\omega t + \vec{k}\cdot\vec{x}_i)}\hat{e}_z \times \partial_t \vec{\xi}_i(t) = \sum_{i=1}^{\mathcal{N}} \int_{-\infty}^{\infty} dt\,(-1)(-i\omega)e^{-i(\omega t + \vec{k}\cdot\vec{x}_i)}\hat{e}_z \times \xi_i(t) + O(\xi^2)$$

$$= +i\omega\,\hat{e}_z \times \vec{\xi}(\vec{k}, \omega) \ .$$

191

In the first step, the first minus sign is from moving the time derivative to act on the exponential using integration by parts. The Fourier transformed equation of motion now takes the simple form

$$\left\{ -\frac{\omega^2}{g^2} + 2J\left[\sum_{\mu=1}^{d} \cos(k_\mu a) - d \right] \right\} \vec{\xi}(\vec{k}, \omega) = -\frac{1}{2}i\omega\,\hat{e}_z \times \vec{\xi}(\vec{k}, \omega) \ .$$

Using $\hat{e}_z \times \hat{e}_x = +\hat{e}_y$ and $\hat{e}_z \times \hat{e}_y = -\hat{e}_x$ and moving everything to the left-hand side yields

$$\begin{pmatrix} -\frac{\omega^2}{g^2} + h(\vec{k}) & -\frac{1}{2}i\omega \\ +\frac{1}{2}i\omega & -\frac{\omega^2}{g^2} + h(\vec{k}) \end{pmatrix} \begin{pmatrix} \xi^x(\vec{k}, \omega) \\ \xi^y(\vec{k}, \omega) \end{pmatrix} = 0$$

where

$$h(\vec{k}) \equiv 2|J|\left[d - \sum_{\mu=1}^{d} \cos(k_\mu a) \right]$$

$$\approx 2|J|\left[d - \sum_{\mu=1}^{d}\left(1 - \frac{1}{2}k_\mu^2 a^2 \right) \right]$$

$$= 2|J|\left[d - \left(d - \frac{1}{2}|\vec{k}|^2 a^2 \right) \right]$$

$$= +|J|a^2 k^2$$

where $k \equiv |\vec{k}|$. On p. 346 of the book, the function $h(k)$ is defined as $h(k) \equiv 4J[2 - \sum_{\mu=1}^{2} \cos(k_\mu a)] \approx 2|J|a^2 k^2$, which is a factor of two larger than the one we have derived. This can be traced back to when we included a factor of $\frac{1}{2}$ in the explicit form of the sum over nearest neighbors. Had we not included this $\frac{1}{2}$, we would be counting each nearest neighbor twice rather than once. In the reference, Wen and Zee have implicitly normalized the coupling J such that the nearest neighbor sum does not include the $\frac{1}{2}$. Thus we can take the results in the text and rescale $J \to \frac{1}{2}J$ in order to match our treatment.

For the antiferromagnet $(J > 0)$, the ground state is where the spins alternate in orientation, so we expand around the ground state as $\vec{n}_i(t) = (-1)^i\hat{e}_z + \vec{\xi}_i(t)$. Let $\vec{Q} \equiv (\frac{\pi}{a}, \frac{\pi}{a})$ be a 2-dimensional vector. We can write $(-1)^i = e^{i\vec{Q}\cdot\vec{x}_i}$, where $\vec{x}_i = x_i^\mu\hat{e}_\mu$ is the coordinate of the

i^{th} lattice site. Thus, for the antiferromagnet we have:

$$\sum_{j=1}^{\mathcal{N}} \mathcal{J}_{ij}(-1)^j = \sum_{j=1}^{\mathcal{N}} \left[J \sum_{\mu=1}^{d} \left(\delta_{i,j+\hat{e}_\mu} + \delta_{i,j-\hat{e}_\mu} \right) - 2\nu\, \delta_{ij} \right] e^{i\vec{Q}\cdot\vec{x}_j}$$

$$= J \sum_{\mu=1}^{d} \left(e^{i\vec{Q}\cdot(\vec{x}_i - a\hat{e}_\mu)} + e^{i\vec{Q}\cdot(\vec{x}_i + a\hat{e}_\mu)} \right) - 2\nu\, e^{i\vec{Q}\cdot\vec{x}_i}$$

$$= 2\, e^{i\vec{Q}\cdot\vec{x}_i} \left[J \sum_{\mu=1}^{d} \cos(Q_\mu a) - \nu \right]$$

$$= 2\, e^{i\vec{Q}\cdot\vec{x}_i} \left[J \sum_{\mu=1}^{d} \cos\pi - \nu \right]$$

$$= 2\, e^{i\vec{Q}\cdot\vec{x}_i}(-1)\,(Jd + \nu)\ .$$

Therefore the z-component of the equation of motion for the antiferromagnet fixes the Lagrange multiplier to be

$$\nu = -Jd$$

rather than $+Jd$. The matrix $\mathcal{J}_{ij}$ for the antiferromagnet is therefore:

$$\mathcal{J}_{ij} = J \left[\sum_{\mu=1}^{d} \left(\delta_{i,j+\hat{e}_\mu} + \delta_{i,j-\hat{e}_\mu} \right) + 2d\, \delta_{ij} \right]\ .$$

For the antiferromagnet, the Berry's phase term becomes:

$$\sum_{i=1}^{\mathcal{N}} \int_{-\infty}^{\infty} dt\, e^{-i(\omega t + \vec{k}\cdot\vec{x}_i)} \vec{\xi}_i(t) \times \partial_t \vec{\xi}_i(t) = +i\omega\, \hat{e}_z \times \vec{\xi}(\vec{k} + \vec{Q}, \omega)$$

again dropping terms of order ξ^2. This results in the equations of motion

$$\left[-\frac{\omega^2}{g^2} + f(\vec{k}) \right] \xi^x(\vec{k}, \omega) = +\frac{1}{2} i\omega\, \xi^y(\vec{k} + \vec{Q}, \omega)$$

$$\left[-\frac{\omega^2}{g^2} + f(\vec{k}) \right] \xi^y(\vec{k}, \omega) = -\frac{1}{2} i\omega\, \xi^x(\vec{k} + \vec{Q}, \omega)\ .$$

where $f(\vec{k}) \equiv 2J \left[d + \sum_{\mu=1}^{d} \cos(k_\mu a) \right] = h(\vec{k} + \vec{Q})$.

To put this into a matrix notation, we need to shift the second equation by $\vec{k} \to \vec{k} + \vec{Q}$ and use $\vec{\xi}(\vec{k} + 2\vec{Q}, \omega) = \vec{\xi}(\vec{k}, \omega)$. We obtain

$$\begin{pmatrix} -\frac{\omega^2}{g^2} + f(\vec{k}) & -\frac{1}{2} i\omega \\ +\frac{1}{2} i\omega & -\frac{\omega^2}{g^2} + f(\vec{k} + \vec{Q}) \end{pmatrix} \begin{pmatrix} \xi^x(\vec{k}, \omega) \\ \xi^y(\vec{k} + \vec{Q}, \omega) \end{pmatrix} = 0\ .$$

Note the argument $\vec{k} + \vec{Q}$ in the lower-right component of the above matrix. This is accidentally absent in the book, but present in equation (7) of the reference.

VI.6   Surface Growth and Field Theory

3. Field theory can often be cast into apparently rather different forms by a change of variable. Show that by writing $U = e^{\frac{1}{2}gh}$ we can change the action (7) to

$$S = \frac{2}{g^2} \int d^D x \, dt \left(U^{-1} \frac{\partial}{\partial t} U - U^{-1} \nabla^2 U \right)^2$$

a kind of nonlinear σ model.

$$(7) \qquad S = \frac{1}{2} \int d^D x \, dt \left[\left(\frac{\partial}{\partial t} - \nabla^2 \right) h - \frac{g}{2} (\nabla h)^2 \right]^2$$

Solution:

Let $h = \frac{2}{g} \ln U$. Then $\nabla_j \ln U = U^{-1} \nabla_j U \implies \nabla_i \nabla_j \ln U = -U^{-2} \nabla_i U \nabla_j U + U^{-1} \nabla_i \nabla_j U$ implies

$$\nabla^2 h + \frac{g}{2} (\nabla h)^2 = -\frac{2}{g} \left(-U^{-2} \nabla_i U \nabla^i U + U^{-1} \nabla^2 U \right) + \frac{g}{2} \left(\frac{2}{g} \right)^2 (U^{-1} \nabla_i U)(U^{-1} \nabla^i U)$$

$$= \frac{2}{g} U^{-1} \nabla^2 U$$

Since $\partial_t \ln U = U^{-1} \partial_t U$, we have

$$S = \frac{2}{g} \int d^D x \, dt \left[U^{-1} \left(\partial_t - \nabla^2 \right) U \right]^2 .$$

VI.7   Disorder: Replicas and Grassmannian Symmetry

1. Work out the field theory that will allow you to study Anderson localization. [Hint: Consider the object

$$\left\langle \left(\frac{1}{z - H} \right) (x, y) \left(\frac{1}{w - H} \right) (y, x) \right\rangle$$

for two complex numbers z and w. You will have to introduce two sets of replica fields, commonly denoted by φ_a^+ and φ_a^-.]

Solution:

Using the identities in the chapter we can write

$$\left(\frac{1}{z - H} \right) (x, y) = i \lim_{n \to 0} \int \mathcal{D}\varphi_+^\dagger \mathcal{D}\varphi_+ \, e^{i \int d^D x' \{ \partial \vec{\varphi}_+^\dagger \partial \vec{\varphi}_+ + [V(x') - z] \vec{\varphi}_+^\dagger \vec{\varphi}_+ \}} \varphi_{+1}(x) \varphi_{+1}^\dagger (y)$$

where we have denoted the n-dimensional vector of replica fields by

$$\vec{\varphi}_+ = \begin{pmatrix} \varphi_{+1} \\ \vdots \\ \varphi_{+n} \end{pmatrix} .$$

194

Similarly, we can also write

$$\left(\frac{1}{w-H}\right)(y,x) = i \lim_{m\to 0} \int \mathcal{D}\varphi_-^\dagger \mathcal{D}\varphi_- \, e^{i\int d^D x' \{\partial\vec{\varphi}_-^\dagger \partial\vec{\varphi}_- + [V(x')-w]\vec{\varphi}_-^\dagger \vec{\varphi}_-\}} \varphi_{-1}(y)\varphi_{-1}^\dagger(x)$$

where we have denoted the additional m-dimensional vector of replica fields by

$$\vec{\varphi}_- = \begin{pmatrix} \varphi_{-1} \\ \vdots \\ \varphi_{-m} \end{pmatrix}.$$

Multiplying these together, we obtain

$$\left(\frac{1}{z-H}\right)(x,y)\left(\frac{1}{w-H}\right)(y,x) =$$

$$i^2 \lim_{n,m\to 0} \int \mathcal{D}\varphi_+^\dagger \mathcal{D}\varphi_+ \mathcal{D}\varphi_-^\dagger \mathcal{D}\varphi_- \, e^{i\int d^D x'\{\partial\vec{\varphi}_+^\dagger \partial\vec{\varphi}_+ + \partial\vec{\varphi}_-^\dagger \partial\vec{\varphi}_- - z\vec{\varphi}_+^\dagger \vec{\varphi}_+ - w\vec{\varphi}_-^\dagger \vec{\varphi}_-\}} \times$$

$$e^{i\int d^D x' \, V(x')(\vec{\varphi}_+^\dagger \vec{\varphi}_+ + \vec{\varphi}_-^\dagger \vec{\varphi}_-)} \, \varphi_{+1}(x)\varphi_{+1}^\dagger(y)\varphi_{-1}(y)\varphi_{-1}^\dagger(x)$$

where we have separated out the terms linear in $V(x)$ to prepare for averaging over disorder.

We now carry out the average $\langle \mathcal{O}(V)\rangle = \mathcal{N}\int \mathcal{D}V \, e^{-\frac{1}{2g^2}\int d^D x \, V(x)^2} \mathcal{O}(V)$ using the formula

$$\mathcal{N}\int \mathcal{D}V \, e^{-\frac{1}{2}V\cdot M\cdot V + j\cdot V} = e^{\frac{1}{2}j\cdot M^{-1}\cdot j}$$

where in our case $j = i(\vec{\varphi}_+^\dagger \vec{\varphi}_+ + \vec{\varphi}_-^\dagger \vec{\varphi}_-)$ and $M = 1/g^2$. We arrive at the expression

$$\left\langle \left(\frac{1}{z-H}\right)(x,y)\left(\frac{1}{w-H}\right)(y,x)\right\rangle =$$

$$-\lim_{n,m\to 0}\int \mathcal{D}\varphi_+^\dagger \mathcal{D}\varphi_+ \mathcal{D}\varphi_-^\dagger \mathcal{D}\varphi_- \, e^{i\int d^D x' \mathcal{L}(x')} \, \varphi_{+1}(x)\varphi_{+1}^\dagger(y)\varphi_{-1}(y)\varphi_{-1}^\dagger(x)$$

with the Lagrangian density

$$\mathcal{L} = \partial\vec{\varphi}_+^\dagger \partial\vec{\varphi}_+ + \partial\vec{\varphi}_-^\dagger \partial\vec{\varphi}_- - z\,\vec{\varphi}_+^\dagger \vec{\varphi}_+ - w\,\vec{\varphi}_-^\dagger \vec{\varphi}_- + i\frac{1}{2}g^2\left(\vec{\varphi}_+^\dagger \vec{\varphi}_+ + \vec{\varphi}_-^\dagger \vec{\varphi}_-\right)^2.$$

2. As another example from the literature on disorder, consider the following problem. Place N points randomly in a D-dimensional Euclidean space of volume V. Denote the locations of the points by $\vec{x}_i$ $(i = 1, ..., N)$. Let

$$f(\vec{x}) = -\int \frac{d^D k}{(2\pi)^D} \frac{e^{i\vec{k}\cdot\vec{x}}}{\vec{k}^2 + m^2}$$

Consider the N by N matrix $H_{ij} = f(\vec{x}_i - \vec{x}_j)$. Calculate $\rho(E)$, the density of eigenvalues of H as we average over the ensemble of matrices, in the limit $N \to \infty, V \to \infty$, with the density of points $\rho_0 \equiv N/V$ held fixed. Hint: Use the replica method and arrive at the field theory action

$$S = \int d^D x \left[\sum_{a=1}^{n} \left(|\nabla\varphi_a|^2 + m^2|\varphi_a|^2 \right) - \rho_0\, e^{-(1/z)\sum_{a=1}^n |\varphi_a|^2} \right]$$

This problem is not entirely trivial; if you need help consult M. Mèzard et al., *Nucl. Phy.* B559: 689, 2000, cond-mat/9906135.

Solution:

The same steps leading to equation (3) on p. 352 result in

$$\mathrm{tr}\left(\frac{1}{z-H}\right) = i \lim_{n\to 0} \sum_{j=1}^{N} \int \left[\prod_{a=1}^{n} \prod_{i=1}^{N} d\varphi_{ai}^\dagger d\varphi_{ai} \right] e^{i\sum_{a=1}^{n}\sum_{i,j=1}^{N} \varphi_{ai}^\dagger (H_{ij}-z\delta_{ij})\varphi_{aj}} \varphi_{j1}\varphi_{j1}^\dagger .$$

Then to "average over disorder," we average over the locations of the N points $\{\vec{x}_i\}_{i=1}^N$:

$$\langle(...)\rangle = \int \left[\prod_{i=1}^{N} \frac{d^D x_i}{V} \right] (...) .$$

Just as we find the propagator by taking derivatives of the partition function, to find the operator $\mathrm{tr}[1/(z-H)]$ we will instead compute the function

$$\xi_N \equiv \int \left[\prod_{i=1}^{N} \frac{d^D x_i}{V} \right] \int d^{nN}\varphi^\dagger d^{nN}\varphi\, e^{i\sum_{a=1}^{n}\sum_{i,j=1}^{N} \varphi_{ai}^\dagger (H_{ij}-z\delta_{ij})\varphi_{aj}}$$

where $d^{nN}\varphi^\dagger d^{nN}\varphi \equiv \prod_{a=1}^{n}\prod_{i=1}^{N} d\varphi_{ai}^\dagger d\varphi_{ai}$. If we define the functions

$$\phi_a(\vec{x}) \equiv \sum_{i=1}^{N} \delta^D(\vec{x} - \vec{x}_i)\varphi_{ai}$$

196

and their hermitian conjugates, then

$$\int d^D x \, d^D y \, \phi_a^\dagger(\vec{x}) f(\vec{x} - \vec{y}) \phi_a(\vec{y}) =$$

$$\int d^D x \, d^D y \left(\sum_{i=1}^{N} \delta^D(\vec{x} - \vec{x}_i) \varphi_{ai}^\dagger \right) f(\vec{x} - \vec{y}) \left(\sum_{j=1}^{N} \delta^D(\vec{y} - \vec{x}_j) \varphi_{aj} \right)$$

$$= \sum_{i,j=1}^{N} \varphi_{ai}^\dagger f(\vec{x}_i - \vec{x}_j) \varphi_{aj} = \sum_{i,j=1}^{N} \varphi_{ai}^\dagger H_{ij} \varphi_{aj} \, ,$$

so we can replace $\sum_{ij} \varphi_{ai}^\dagger H_{ij} \varphi_{aj}$ with $\int d^D x \, d^D y \, \phi_a^\dagger(\vec{x}) f(\vec{x} - \vec{y}) \phi_a(\vec{y})$. To make this a useful substitution, we need to be able to treat the variables $\phi_a(\vec{x})$ and $\phi_a^\dagger(\vec{x})$ as unconstrained field variables. In other words, we want to integrate $\int \mathcal{D}\phi^\dagger \mathcal{D}\phi$ as if $\phi_a(\vec{x})$ and $\phi_a^\dagger(\vec{x})$ were usual bosonic quantum fields that transform as a vector under an $O(n)$ symmetry, and we can do so as long as we include the constraint that $\phi_a(\vec{x}) - \sum_{i=1}^{N} \delta^D(\vec{x} - \vec{x}_i) \varphi_{ai} = 0$ and $\phi_a^\dagger(\vec{x}) - \sum_{i=1}^{N} \delta^D(\vec{x} - \vec{x}_i) \varphi_{ai}^\dagger = 0$. To implement this, insert the number 1 into the path integral as follows:

$$1 = \int \mathcal{D}\phi \, \delta^{(n)}\left[\vec{\phi}(\vec{x}) - \sum_{i=1}^{N} \vec{\varphi}_i \delta^{(D)}(\vec{x} - \vec{x}_i) \right] = \int \mathcal{D}\phi \int \mathcal{D}\mu^\dagger \, e^{i \int d^D x \, \vec{\mu}^\dagger(\vec{x}) \cdot \left[\vec{\phi}(\vec{x}) - \sum_{i=1}^{N} \vec{\varphi}_i \delta^{(D)}(\vec{x} - \vec{x}_i) \right]}$$

and the same for the hermitian conjugate:

$$1 = \int \mathcal{D}\phi^\dagger \, \delta^{(n)}\left[\vec{\phi}^\dagger(\vec{x}) - \sum_{i=1}^{N} \vec{\varphi}_i^\dagger \delta^{(D)}(\vec{x} - \vec{x}_i) \right] = \int \mathcal{D}\phi^\dagger \int \mathcal{D}\mu \, e^{-i \int d^D x \, \vec{\mu}(\vec{x}) \cdot \left[\vec{\phi}^\dagger(\vec{x}) - \sum_{i=1}^{N} \vec{\varphi}_i^\dagger \delta^{(D)}(\vec{x} - \vec{x}_i) \right]}$$

We chose the minus sign in this version of 1 for later convenience. If you are getting bogged down in functional integral notation, then orient yourself with the familiar example from ordinary 1-dimensional calculus:

$$1 = \int_{-\infty}^{\infty} dx \, \delta(x - c) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} \frac{d\mu}{2\pi} \, e^{\pm i \mu (x - c)}$$

where c is some real number. In the functional integrals, the factors of 2π are swept into the definitions of the measures. Inserting both of these factors of 1 into the integral for ξ_N, we arrive at

$$\xi_N =$$

$$\int \left[\prod_{i=1}^{N} \frac{d^D x_i}{V} \right] \int \mathcal{D}\phi^\dagger \mathcal{D}\phi \mathcal{D}\mu^\dagger \mathcal{D}\mu \int d^{nN}\varphi^\dagger d^{nN}\varphi \, \exp\left[i \int d^D x \, d^D y \sum_{a=1}^{n} \phi_a^\dagger(\vec{x}) f(\vec{x} - \vec{y}) \phi_a(\vec{y}) \right] \times$$

$$\exp\left[-i \sum_{a=1}^{n} \sum_{i=1}^{N} z \, \varphi_{ai}^\dagger \varphi_{ai} + \int d^D x \sum_{a=1}^{n} \left\{ i \, \mu_a^\dagger(\vec{x}) [\phi_a(\vec{x}) - \sum_{i=1}^{N} \varphi_{ai} \delta^D(\vec{x} - \vec{x}_i)] + h.c. \right\} \right] .$$

We can now perform the Gaussian integral over the φ_{ai} and $\varphi_{ai}^\dagger$ variables. First note that $\int d^D x\, \mu_a(\vec{x})\delta^D(\vec{x} - \vec{x}_i) = \mu_a(\vec{x}_i)$, and then define the "sources" $J_a(\vec{x}) \equiv -i\mu_a(\vec{x})$. Then the integral we have to perform is

$$I \equiv \int d^{nN}\varphi^\dagger d^{nN}\varphi \, \exp\left\{\sum_{a=1}^{n}\sum_{i=1}^{N}\left[-iz\,\varphi_{ai}^\dagger\varphi_{ai} + i\,\mu_a^\dagger(\vec{x}_i)\varphi_{ai} - i\,\varphi_{ai}^\dagger\mu_a(\vec{x}_i)\right]\right\}$$

$$= \int d^{nN}\varphi^\dagger d^{nN}\varphi \, \exp\left\{\sum_{a=1}^{n}\sum_{i=1}^{N}\left[-iz\,\varphi_{ai}^\dagger\varphi_{ai} + J_a^\dagger(\vec{x}_i)\varphi_{ai} + \varphi_{ai}^\dagger J_a(\vec{x}_i)\right]\right\}$$

$$= \prod_{a=1}^{n}\prod_{i=1}^{N}\int d\varphi_{ai}^\dagger d\varphi_{ai} \, \exp\left\{-(iz)\varphi_{ai}^\dagger\varphi_{ai} + J_a^\dagger(\vec{x}_i)\varphi_{ai} + \varphi_{ai}^\dagger J_a(\vec{x}_i)\right\} \ .$$

Using

$$\int d\varphi^\dagger d\varphi\, e^{-\alpha\,\varphi^\dagger\varphi + J^\dagger\varphi + \varphi^\dagger J} = \frac{2\pi i}{\alpha}\, e^{+\frac{1}{\alpha}J^\dagger J}$$

with $\alpha = iz$, we have

$$I = \prod_{a=1}^{n}\prod_{i=1}^{N}\frac{2\pi}{z}\,\exp\left\{\frac{1}{iz}J_a^\dagger(\vec{x}_i)J_a(\vec{x}_i)\right\}$$

$$= \left(\frac{2\pi}{z}\right)^{nN}\exp\left\{-\frac{i}{z}\sum_{a=1}^{n}\sum_{i=1}^{N}J_a^\dagger(\vec{x}_i)J_a(\vec{x}_i)\right\}$$

$$= \left(\frac{2\pi}{z}\right)^{nN}\exp\left\{-\frac{i}{z}\sum_{a=1}^{n}\sum_{i=1}^{N}\mu_a^\dagger(\vec{x}_i)\mu_a(\vec{x}_i)\right\} \ .$$

Therefore the original integral ξ_N we wanted is now

$$\xi_N =$$
$$\left(\frac{2\pi}{z}\right)^{nN}\int\left[\prod_{i=1}^{N}\frac{d^D x_i}{V}\right]\int \mathcal{D}\phi^\dagger\mathcal{D}\phi\mathcal{D}\mu^\dagger\mathcal{D}\mu\,\exp\left[i\int d^D x\, d^D y\sum_{a=1}^{n}\phi_a^\dagger(\vec{x})f(\vec{x}-\vec{y})\phi_a(\vec{y})\right]\times$$

$$\exp\left\{-\frac{i}{z}\sum_{a=1}^{n}\sum_{i=1}^{N}\mu_a^\dagger(\vec{x}_i)\mu_a(\vec{x}_i) + i\int d^D x\sum_{a=1}^{n}\left[\mu_a^\dagger(\vec{x})\phi_a(\vec{x}) - \phi_a^\dagger(\vec{x})\mu_a(\vec{x})\right]\right\} \ .$$

Now we can perform the Gaussian integral over the fields $\phi_a(\vec{x})$ and $\phi_a^\dagger(\vec{x})$:

$$\int \mathcal{D}\phi^\dagger\mathcal{D}\phi\, e^{i\int d^D x\, d^D y\sum_{a=1}^{n}\phi_a^\dagger(\vec{x})f(\vec{x}-\vec{y})\phi_a(\vec{y}) + \int d^D x\sum_{a=1}^{n}[J_a^\dagger(\vec{x})\phi_a(\vec{x}) + \phi_a^\dagger(\vec{x})J_a(\vec{x})]}$$

$$= (\det f)^{-1}e^{+i\int d^D x\, d^D y\sum_{a=1}^{n}J_a^\dagger(\vec{x})f^{-1}(\vec{x}-\vec{y})J_a(\vec{y})} \ ,$$

where again $J_a(\vec{x}) = -i\mu_a(\vec{x})$. Since $f(x) = -\int\frac{d^D k}{(2\pi)^D}\frac{1}{k^2+m^2}e^{ikx}$, the inverse of $f(x)$ is

$$f^{-1}(x) = -\int\frac{d^D k}{(2\pi)^D}(k^2 + m^2)e^{ikx} \ .$$

198

To convince yourself this is right, compute

$$\int d^D z \, f^{-1}(x-z) f(z-y)$$

$$= (-1)^2 \int \frac{d^D k}{(2\pi)^D} \int \frac{d^D p}{(2\pi)^D} (k^2 + m^2) \frac{1}{p^2 + m^2} \, e^{+ikx} \, e^{-ipy} \int d^D z \, e^{+i(p-k)z}$$

$$= \int \frac{d^D k}{(2\pi)^D} \int \frac{d^D p}{(2\pi)^D} (k^2 + m^2) \frac{1}{p^2 + m^2} \, e^{+ikx} \, e^{-ipy} (2\pi)^D \delta^D(p-k)$$

$$= \int \frac{d^D k}{(2\pi)^D} \, e^{+ik(x-y)} = \delta^D(x-y) . \checkmark$$

To summarize our progress so far, we have

$$\xi_N = \left(\frac{2\pi}{z}\right)^{nN} \int \left[\prod_{i=1}^{N} \frac{d^D x_i}{V}\right] \int \mathcal{D}\mu^\dagger \mathcal{D}\mu \, \exp\left[-\frac{i}{z} \sum_{a=1}^{n} \sum_{i=1}^{N} \mu_a^\dagger(\vec{x}_i) \mu_a(\vec{x}_i)\right] \times$$

$$\exp\left[-\operatorname{tr} \ln f + i \int d^D x \, d^D y \sum_{a=1}^{n} \mu_a^\dagger(\vec{x}) f^{-1}(\vec{x}-\vec{y}) \mu_a(\vec{y})\right] .$$

Given the explicit form of $f^{-1}(\vec{x})$, the last term simplifies (suppress the index a for clarity):

$$\int d^D x \, d^D y \, \mu^\dagger(\vec{x}) f^{-1}(\vec{x}-\vec{y}) \mu(\vec{y}) = \int d^D x \, d^D y \, \mu^\dagger(\vec{x}) \left(-\int \frac{d^D k}{(2\pi)^D} (|\vec{k}|^2 + m^2) \, e^{i\vec{k}\cdot(\vec{x}-\vec{y})}\right) \mu(\vec{y})$$

$$= -\int d^D x \, d^D y \int \frac{d^D k}{(2\pi)^D} \mu^\dagger(\vec{x}) \, e^{+i\vec{k}\cdot\vec{x}} (|\vec{k}|^2 + m^2) \, e^{-i\vec{k}\cdot\vec{y}} \mu(\vec{y})$$

$$= -\int d^D x \, d^D y \int \frac{d^D k}{(2\pi)^D} \mu^\dagger(\vec{x}) \, e^{+i\vec{k}\cdot\vec{x}} \left[(-\nabla_y^2 + m^2) e^{-i\vec{k}\cdot\vec{y}}\right] \mu(\vec{y})$$

$$= -(-1)^2 \int d^D x \, d^D y \left(\int \frac{d^D k}{(2\pi)^D} \, e^{+i\vec{k}\cdot(\vec{x}-\vec{y})}\right) \mu^\dagger(\vec{x})(-\nabla_y^2 + m^2) \mu(\vec{y})$$

$$= -\int d^D x \, d^D y \, \delta^D(\vec{x}-\vec{y}) \mu^\dagger(\vec{x})(-\nabla_y^2 + m^2) \mu(\vec{y})$$

$$= -\int d^D x \, \mu^\dagger(\vec{x})(-\nabla^2 + m^2) \mu(\vec{x})$$

$$= -\int d^D x \left[+\vec{\nabla}\mu^\dagger(\vec{x}) \cdot \vec{\nabla}\mu(\vec{x}) + m^2 \mu^\dagger(\vec{x}) \mu(\vec{x})\right] .$$

Now this is starting to approach the desired result. At this point we have the field theory

$$\xi_N = C \int \mathcal{D}\mu^\dagger \mathcal{D}\mu \, A^N \, e^{-S_0[\mu^\dagger, \mu]}$$

with the action

$$S_0[\mu^\dagger, \mu] = i \int d^D x \sum_{a=1}^{n} \left[\vec{\nabla}\mu_a^\dagger(\vec{x}) \cdot \vec{\nabla}\mu_a(\vec{x}) + m^2 \mu_a^\dagger(\vec{x}) \mu_a(\vec{x})\right] ,$$

199

the function
$$A = \left(\frac{2\pi}{z}\right)^n \int \frac{d^D x}{V} e^{-(i/z)\sum_{a=1}^n \mu_a^\dagger(\vec{x})\mu_a(\vec{x})} \ ,$$

and the irrelevant overall constant
$$C = e^{-\operatorname{tr} \ln f} \ .$$

We will now drop C since it will drop out of all correlation functions as do all overall constants in the path integral. We will also take $n \to 0$ in the prefactor $(2\pi/z)^n$ in A. Passing to a grand canonical formulation of the disorder[16], we compute

$$\xi \equiv \sum_{N=0}^\infty \frac{1}{N!}\xi_N \alpha^N = \int \mathcal{D}\mu^\dagger \mathcal{D}\mu \left(\sum_{N=0}^\infty \frac{1}{N!}(A\alpha)^N\right) e^{-S_0[\mu^\dagger, \mu]} = \int \mathcal{D}\mu^\dagger \mathcal{D}\mu \, e^{A\alpha - S_0[\mu^\dagger, \mu]} \ .$$

We therefore arrive at the result $\xi = \int \mathcal{D}\mu^\dagger \mathcal{D}\mu \, e^{-S}$, where the full action $S = S_0 - \alpha A$ is

$$S[\mu^\dagger, \mu] = \int d^D x \left[i\sum_{a=1}^n \left(\vec{\nabla}\mu_a^\dagger(\vec{x})\cdot\vec{\nabla}\mu_a(\vec{x}) + m^2\mu_a^\dagger(\vec{x})\mu_a(\vec{x}) \right) - \frac{\alpha}{V}\exp\left(-\frac{i}{z}\sum_{a=1}^n \mu_a^\dagger(\vec{x})\mu_a(\vec{x})\right) \right] \ .$$

The last thing to determine is the meaning of the parameter α. The average number of points is[17] $N = \alpha\langle A\rangle$, so that in the limit $n \to 0$ we have $N = \alpha$ and hence $\alpha/V = \rho_0$, the density of points. We have arrived at the action

$$S[\mu^\dagger, \mu] = \int d^D x \left[i\sum_{a=1}^n \left(\vec{\nabla}\mu_a^\dagger(\vec{x})\cdot\vec{\nabla}\mu_a(\vec{x}) + m^2\mu_a^\dagger(\vec{x})\mu_a(\vec{x}) \right) - \rho_0\exp\left(-\frac{i}{z}\sum_{a=1}^n \mu_a^\dagger(\vec{x})\mu_a(\vec{x})\right) \right] \ .$$

Perform the field redefinition $\mu \to -i\mu$ while keeping $\mu^\dagger$ unchanged to obtain

$$S[\mu^\dagger, \mu] = \int d^D x \left[\sum_{a=1}^n \left(\vec{\nabla}\mu_a^\dagger(\vec{x})\cdot\vec{\nabla}\mu_a(\vec{x}) + m^2\mu_a^\dagger(\vec{x})\mu_a(\vec{x}) \right) - \rho_0\exp\left(-\frac{1}{z}\sum_{a=1}^n \mu_a^\dagger(\vec{x})\mu_a(\vec{x})\right) \right]$$

which is the desired result.

---

[16]This is just as in statistical mechanics. The partition function $Z = \operatorname{tr} e^{-\beta H}$ can be put into a grand canonical formulation by introducing a chemical potential μ via $Z(\mu) = \operatorname{tr} \sum_{N=0}^\infty \frac{1}{N!}e^{-\beta(H-\mu N)}$.

[17]The average particle number in a grand canonical ensemble described by the partition function $Z(\mu)$ is $\langle N\rangle = \frac{1}{\beta}\frac{\partial}{\partial\mu}\ln Z(\mu)$. In our case, we have $\beta = 1$ and $\mu = \ln\alpha$. The average value of the function A defined previously is $\langle A\rangle = \frac{\partial}{\partial\alpha}\ln\xi$, and hence $\alpha\langle A\rangle = \alpha\frac{\partial}{\partial\alpha}\ln\xi = \frac{\partial}{\partial\ln\alpha}\ln\xi$. Since $\ln\alpha = \mu$, we have $\alpha\langle A\rangle = \frac{\partial}{\partial\mu}\ln\xi = N$.

VI.8   Renormalization Group Flow as a Natural Concept in High Energy and Condensed Matter Physics

1. Show that the solution of $dg/dt = -bg^3 + ...$ is given by

$$\frac{1}{\alpha(t)} = \frac{1}{\alpha(0)} + 8\pi bt + ...$$

where we defined $\alpha(t) = g(t)^2/4\pi$.

Solution:

$$\int_{g_0}^{g} \frac{dg'}{g'^3} = -b \int_{0}^{t} dt' \implies -\frac{1}{2}\left(\frac{1}{g^2} - \frac{1}{g_0^2}\right) = -b\,t \implies \frac{1}{g^2} = \frac{1}{g_0^2} + 2bt$$

Define $g^2 \equiv 4\pi\alpha$ and multiply by 4π to get

$$\frac{1}{\alpha} = \frac{1}{\alpha_0} + 8\pi b\,t \ .$$

2. In our discussion of the renormalization group, in $\lambda\varphi^4$ or in QED, for the sake of simplicity we assumed that the mass m of the particle is much smaller than μ and thus set m equal to zero. But nothing in the renormalization group idea tells us that we can't flow to a mass scale below m. Indeed, in particle physics many orders of magnitude separate the top quark mass m_t from the up quark mass m_u. We might want to study how the strong interaction coupling flows from some mass scale far above m_t down to some mass scale μ below m_t but still large compared to m_u. As a crude approximation, people often set any mass m below μ equal to zero and any m above μ to infinity (i.e., not contributing to the renormalization group flow). In reality, as μ approaches m from above the particle starts to contribute less and drops out as μ becomes much less than m. Taking either the $\lambda\varphi^4$ theory or QED study this so-called threshold effect.

Solution:

First consider QED. Equation (6) on p. 358 gives the renormalization group flow equation for the QED coupling:

$$\mu\frac{d}{d\mu}e(\mu) = -\frac{1}{2}e(\mu)^3\,\mu\frac{d}{d\mu}\Pi(\mu^2) + O(e^5)$$

The lowest order solution to this is

$$\frac{1}{e^2(\mu)} = \frac{1}{e^2(\mu_0)} - \frac{1}{6\pi^2}\ln\frac{\mu}{\mu_0} \ .$$

Suppose we have $n+1$ electrons in this theory, one having mass m and the others massless, and take the initial condition to be at some superheavy mass scale M.

For $m < \mu < M$, all $n + 1$ particles contribute to the vacuum polarization function, and we have

$$\frac{1}{e^2(\mu)} = \frac{1}{e^2(M)} + \frac{n+1}{6\pi^2} \ln \frac{M}{\mu} \ .$$

For $0 < \mu < m$, the lowest order approximation is to use

$$\frac{1}{e^2(\mu)} = \frac{1}{e^2(M)} + \frac{n}{6\pi^2} \ln \frac{M}{\mu} \ .$$

Including the threshold correction at the scale m amounts to replacing the lowest order expression with the following:

$$\begin{aligned}
\frac{1}{e^2(\mu)} &= \frac{1}{e^2(m)} + \frac{n}{6\pi^2} \ln \frac{m}{\mu} \\
&= \left(\frac{1}{e^2(M)} + \frac{n+1}{6\pi^2} \ln \frac{M}{m} \right) + \frac{n}{6\pi^2} \ln \frac{m}{\mu} \\
&= \frac{1}{e^2(M)} + \frac{n}{6\pi^2} \ln \frac{M}{\mu} + \frac{1}{6\pi^2} \ln \frac{M}{m} \ .
\end{aligned}$$

We see that there is an extra term that depends explicitly on m. This result is usually written in terms of the μ-dependent coupling for $\mu > m$, which we now denote by[18] $e_G(\mu)$. With $e^{-2}(M) = e_G^{-2}(\mu) - \left(\frac{n+1}{6\pi^2} \right) \ln \frac{M}{\mu}$, we can rewrite the threshold-corrected coupling for $\mu < m$ as:

$$\begin{aligned}
\frac{1}{e^2(\mu)} &= \left[\frac{1}{e_G^2(\mu)} - \left(\frac{n+1}{6\pi^2} \right) \ln \frac{M}{\mu} \right] + \frac{n}{6\pi^2} \ln \frac{M}{\mu} + \frac{1}{6\pi^2} \ln \frac{M}{m} \\
&= \frac{1}{e_G^2(\mu)} - \frac{1}{6\pi^2} \ln \frac{m}{\mu} \ .
\end{aligned}$$

The analysis for φ^4 scalar field theory proceeds in the same way. Equation (5) on p. 357 along with appendix 1 in chapter III.1 (equation (14) on p. 168) gives the renormalization group flow equation for $\lambda \varphi^4$ theory:

$$\mu \frac{d}{d\mu} \lambda(\mu) = -\frac{1}{16} \lambda(\mu)^2 \, \mu \frac{d}{d\mu} \Pi(\mu^2) + O(\lambda^3)$$

where we have defined the function

$$\Pi(s) \equiv \frac{1}{2\pi^2} \int_0^1 dx \, \ln \left(\frac{\Lambda^2}{m^2 - x(1-x)s} \right)$$

---

[18]We use this notation to evoke the connection to threshold corrections in grand unified theories. See chapter VII.6 in the main text, as well as S. Weinberg, "Effective Gauge Theories," Phys. Lett. 91B, No. 1 (1980) and L. Hall, "Grand Unification of Effective Gauge Theories," Nucl. Phys. B178 (1981) 75-124.

using a notation intentionally evocative of the vacuum polarization function in QED. [Here Λ is an arbitrary upper cutoff on the momentum, and m is the mass of the quanta ("mesons") of the scalar field φ.]

The analysis proceeds in exactly the same fashion as for QED: take $n+1$ mesons, one with mass m and the others massless, and assume an initial condition at $M \gg m$.

For $m < \mu < M$, we use
$$\frac{1}{\lambda(\mu)} = \frac{1}{\lambda(M)} + \frac{n+1}{16\pi^2} \ln \frac{M}{\mu} \ .$$

For $0 < \mu < m$ without the threshold correction we use $\lambda^{-1}(\mu) = \lambda^{-1}(M) + \frac{n}{16\pi^2} \ln \frac{M}{\mu}$, but with the threshold correction we use instead

$$\begin{aligned}
\frac{1}{\lambda(\mu)} &= \frac{1}{\lambda(m)} + \frac{n}{16\pi^2} \ln \frac{m}{\mu} \\
&= \left(\frac{1}{\lambda(M)} + \frac{n+1}{16\pi^2} \ln \frac{M}{m} \right) + \frac{n}{16\pi^2} \ln \frac{m}{\mu} \\
&= \frac{1}{\lambda(M)} + \frac{n}{16\pi^2} \ln \frac{M}{\mu} + \frac{1}{16\pi^2} \ln \frac{M}{m} \ .
\end{aligned}$$

As before, we relabel the high-energy coupling as $\lambda_G(\mu)$ and rewrite the low-energy coupling using $\lambda^{-1}(M) = \lambda_G^{-1}(\mu) - \frac{n+1}{16\pi^2} \ln \frac{M}{\mu}$:

$$\begin{aligned}
\frac{1}{\lambda(\mu)} &= \left[\frac{1}{\lambda_G(\mu)} - \frac{n+1}{16\pi^2} \ln \frac{M}{\mu} \right] + \frac{n}{16\pi^2} \ln \frac{M}{\mu} + \frac{1}{16\pi^2} \ln \frac{M}{m} \\
&= \frac{1}{\lambda_G(\mu)} - \frac{1}{16\pi^2} \ln \frac{m}{\mu} \ .
\end{aligned}$$

4. In $\tilde{S}(h)$ only derivatives of the field h can appear and not the field itself. (Since the transformation $h(\vec{x},t) \to h(\vec{x},t) + c$ with c a constant corresponds to a trivial shift of where we measure the surface height from, the physics must be invariant under this transformation.) Terms involving only one power of h cannot appear since they are all total divergences. Thus, $\tilde{S}(h)$ must start with terms quadratic in h. Verify that the $\tilde{S}(h)$ given in (17) is indeed the most general. A term proportional to $(\nabla h)^2$ is also allowed by symmetries and is in fact generated. However, such a term can be eliminated by transforming to a moving coordinate frame $h \to h + ct$.

Solution:

The action is supposed to be the most general compatible with invariance under the Galilean transformation

$$h(\vec{x},t) \to h'(\vec{x},t) = h(\vec{x} + g\,\vec{u}\,t,\, t) + \vec{u}\cdot\vec{x} + \tfrac{1}{2}gu^2 t$$

with $\vec{u}$ a constant velocity. In the solution for VI.8.3 it is shown that the combination

$$\partial_t h - \frac{g}{2}(\vec{\nabla} h)^2$$

is invariant under the Galilean transformation, as is $\nabla^2 h$. Thus in general the action should be a linear combination of these terms:
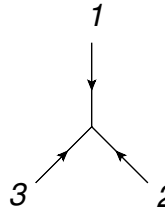
$$X \equiv \alpha\left(\partial_t h - \frac{g}{2}(\vec{\nabla} h)^2\right) - \beta\nabla^2 h \ .$$

As discussed in the problem, the action should also not involve terms linear in h since all such terms are total divergences. Thus to lowest order in h the action is proportional to $\int d^D x\, dt\, X^2$, which is the form given in (17).

6. Calculate the h propagator to one loop order. Extract the coefficients of the ω^2 and k^4 terms in a low frequency and wave number expansion of the inverse propagator and determine α and β.
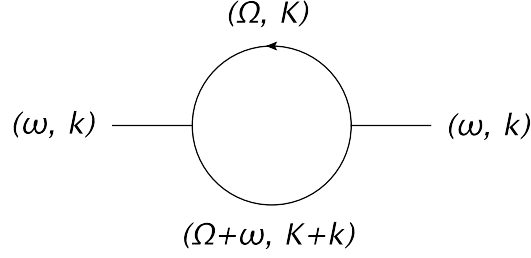
Solution:

We need the free-field propagator and the cubic vertex (all momenta point into the vertex):

$$\frac{\overline{\omega,\ k}}{} \quad = \quad \frac{1}{\omega^2 + k^4}$$

$$= \ (i\omega_1 + k_1{}^2)\, k_2 \bullet k_3 \ + \ (i\omega_2 + k_2{}^2)\, k_3 \bullet k_1 \ + \ (i\omega_3 + k_3{}^2)\, k_1 \bullet k_2$$

Fortunately the quartic vertex does not contribute to this order. The only diagram we need to compute is

Calculating this diagram results in a self-energy

$$\Sigma(\omega, k) = \tfrac{1}{2} \left[\omega^2 + \frac{4d^2 - d - 6}{d(d+2)} k^4 \right] \int_\mu^\Lambda \frac{d^d K}{(2\pi)^d} \frac{1}{K^2} \ .$$

Here μ and Λ are arbitrary lower and upper cutoffs, respectively, on the integral. The idea is to integrate over an infinitesimal shell, so that the lower cutoff can be taken as $\mu = (1 - \delta L)\Lambda$. The integral is now

$$\int_\mu^\Lambda \frac{d^d K}{(2\pi)^d} \frac{1}{K^2} = \frac{\Lambda^{d-2}}{2^{d-1} \pi^{d/2} \Gamma(\frac{d}{2})} \delta L \equiv f(\Lambda, d) \ \delta L \ .$$

Adding the 1-loop self-energy to the zeroth-order propagator, we obtain the desired coefficients

$$\alpha = 1 - \frac{g^2}{8} f(\Lambda, d) \ \delta L$$

$$\beta = 1 - \frac{g^2}{8} \frac{4d^2 - d - 6}{d(d+2)} f(\Lambda, d) \ \delta L \ .$$

See M. Karder and A. Zee, "Matrix generalizations of some dynamic field theories," Nucl. Phys. B464 (1996) 449-462 for further discussion.

VII   Grand Unification

VII.1   Quantizing Yang-Mills Theory and Lattice Gauge Theory

3. Consider a lattice gauge theory in $(D + 1)$-dimensional space with the lattice spacing a in D-dimensional space and b in the extra dimension. Obtain the continuum D-dimensional field theory in the limit $a \to 0$ with b kept fixed.

Solution:

Consider a rectangular plaquette P with sides of unequal lengths a and b:
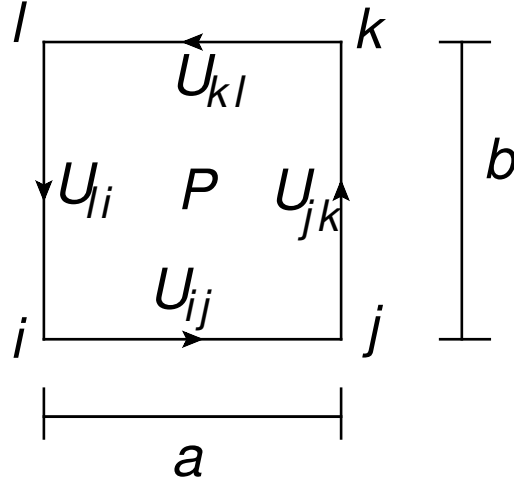
We have $S(P) = \mathrm{Re}\ \mathrm{tr} U_{ij} U_{jk} U_{k\ell} U_{\ell i}$, where

$$U_{ij} = e^{+iaA_\mu(x)} \ , \quad \hat{\mu} = \frac{1}{a}(x_j - x_i) \ , \quad x = \frac{1}{2}(x_i + x_j)$$

$$U_{jk} = e^{+ibA_\nu(y)} \ , \quad \hat{\nu} = \frac{1}{b}(x_k - x_j) \ , \quad y = \frac{1}{2}(x_j + x_k) = x + \frac{a}{2}\hat{\mu} + \frac{b}{2}\hat{\nu}$$

$$U_{k\ell} = e^{-iaA_\mu(x')} \ , \quad x' = x + b\hat{\nu}$$

$$U_{\ell i} = e^{-ibA_\nu(y')} \ , \quad y' = y - a\hat{\mu} = x - \frac{a}{2}\hat{\mu} + \frac{b}{2}\hat{\nu}$$

Since U_{jk} and $U_{\ell i}$ are inverses when evaluated at the point $x + \frac{b}{2}\hat{\nu}$, let $U_{jk}(x + \frac{b}{2}\hat{\nu}) \equiv U$ and $U_{\ell i}(x + \frac{b}{2}\hat{\nu}) = U_{i\ell}^\dagger(x + \frac{b}{2}\hat{\nu}) \equiv U^\dagger$. Taylor expanding in powers of a, we have

- $U_{ij} = e^{+iaA_\mu(x)} = I + iaA_\mu(x) - \frac{1}{2}a^2 A_\mu(x)A_\mu(x) + O(a^3)$

- $U_{jk}(x + \frac{a}{2}\hat{\mu} + \frac{b}{2}\hat{\nu}) = U + \frac{a}{2}\partial_\mu U + \frac{1}{2}\left(\frac{a}{2}\right)^2 \partial_\mu\partial_\mu U + O(a^3)$

- $U_{k\ell} = e^{-iaA_\mu(x')} = I - iaA_\mu(x') - \frac{1}{2}a^2 A_\mu(x')A_\mu(x') + O(a^3)$

- $U_{\ell i}(x - \frac{a}{2}\hat{\mu} + \frac{b}{2}\hat{\nu}) = U^\dagger - \frac{a}{2}\partial_\mu U^\dagger + \frac{1}{2}\left(-\frac{a}{2}\right)^2 \partial_\mu\partial_\mu U^\dagger + O(a^3)$

<center>206</center>

For brevity suppress the direction μ and write $A_\mu(x) \equiv A$ and $A_\mu(x') \equiv A'$. The trace of the product of these four matrices (ignoring terms of $O(a^3)$ and higher) is

$\operatorname{tr} U_{ij} U_{jk} U_{k\ell} U_{\ell i}$

$$= \operatorname{tr} \left[I + iaA - \frac{1}{2}a^2 A^2 \right] \left[U + \frac{a}{2}\partial U + \frac{a^2}{8}\partial^2 U \right] \left[I - iaA' - \frac{1}{2}a^2 A'^2 \right] \left[U^\dagger - \frac{a}{2}\partial U^\dagger + \frac{a^2}{8}\partial^2 U^\dagger \right]$$

$$= \operatorname{tr} \left[U + a\left(\frac{1}{2}\partial U + iAU\right) + \frac{1}{2}a^2\left(\frac{1}{4}\partial^2 U + iA\partial U - A^2 U\right) \right] \times$$
$$\left[U^\dagger - a\left(\frac{1}{2}\partial U^\dagger + iA'U^\dagger\right) + \frac{1}{2}a^2\left(\frac{1}{4}\partial^2 U^\dagger + iA'\partial U^\dagger - A'^2 U^\dagger\right) \right]$$

$$= \operatorname{tr}[\, UU^\dagger + a\left(\frac{1}{2}\partial UU^\dagger + iAUU^\dagger - \frac{1}{2}U\partial U^\dagger - iUA'U^\dagger\right)]$$
$$+ a^2 \operatorname{tr}[\, \frac{1}{2}U\left(\frac{1}{4}\partial^2 U^\dagger + iA'\partial U^\dagger - A'^2 U^\dagger\right) - \left(\frac{1}{2}\partial U + iAU\right)\left(\frac{1}{2}\partial U^\dagger + iA'U^\dagger\right)$$
$$+ \frac{1}{2}\left(\frac{1}{4}\partial^2 U + iA\partial U - A^2 U\right)U^\dagger \,]$$

Since $UU^\dagger = I$, the first term is a constant and can be dropped. Also, $\operatorname{tr}AUU^\dagger = \operatorname{tr}A = 0$. Using the cyclicity of the trace, we also have $\operatorname{tr}UA'U^\dagger = \operatorname{tr}A'U^\dagger U = \operatorname{tr}A' = 0$.

Since $U = e^{ibA_\nu(y)}$, we have $\partial U = Uib\,\partial A_\nu(y)$, and so

$$\operatorname{tr}(\partial UU^\dagger - U\partial U^\dagger) = 2\operatorname{tr}(\partial UU^\dagger) \propto \operatorname{tr}(U\partial AU^\dagger) = \operatorname{tr}(\partial AUU^\dagger) = \operatorname{tr}(\partial A) = 0 \,.$$

Therefore the $O(a)$ terms in $\operatorname{tr} U_{ij}U_{jk}U_{k\ell}U_{\ell i}$ are zero. Dropping the $O(a^0)$ additive constant, we have:

$$\operatorname{tr} U_{ij}U_{jk}U_{k\ell}U_{\ell i}$$

$$= \frac{1}{8}a^2 \operatorname{tr}\left[U\partial^2 U^\dagger + \partial^2 U U^\dagger - 2\partial U \partial U^\dagger\right]$$
$$+ i\frac{1}{2}a^2 \operatorname{tr}\left[U A'\partial U^\dagger + A\partial U U^\dagger - AU\partial U^\dagger - \partial U A'U^\dagger\right]$$
$$+ a^2 \operatorname{tr}\left[-\frac{1}{2}U A'^2 U^\dagger + AU A'U^\dagger - \frac{1}{2}A^2 UU^\dagger\right]$$

$$= -\frac{1}{2}a^2 \operatorname{tr}\left[\partial U \partial U^\dagger\right]$$
$$+ i\frac{1}{2}a^2 \operatorname{tr}\left[(\partial U U^\dagger - U\partial U^\dagger)A + (\partial U^\dagger U - U^\dagger \partial U)A'\right]$$
$$- \frac{1}{2}a^2 \operatorname{tr}\left[A^2 + A'^2 - 2\,AU A'U^\dagger\right]$$

Consider a derivative operator defined as $\mathcal{D}U = \partial U + iAU - iUA'$. This is the covariant derivative appropriate for a field U transforming as $\sim N \otimes \bar{N}$ under $SU(N)$ gauge transformations. To clarify this remark, return to ordinary continuum field theory for a moment and consider a matter field $\psi \sim N \otimes \bar{N}$ of $SU(N)$. This has one lower index $a = 1, ..., N$ and one upper index $\bar{a} = 1, ..., N$, meaning that the components of ψ are $\psi_a{}^{\bar{a}}$. The gauge-covariant derivative of ψ is

$$(\mathcal{D}_\mu \psi)_a{}^{\bar{a}} = \partial_\mu \psi_a{}^{\bar{a}} + i \sum_{I=1}^{N^2-1}\left[A_\mu^I (T_N^I)_a{}^b \psi_b{}^{\bar{a}} + A_\mu^I (T_{\bar{N}}^I)^{\bar{a}}{}_{\bar{b}}\psi_a{}^{\bar{b}}\right]$$

where T_N^I are the generators of the N-representation ("fundamental") of $SU(N)$, and $T_{\bar{N}}^I$ are the generators of the $\bar{N}$-representation ("anti-fundamental") of $SU(N)$. (As usual, the index $I = 1, ..., N^2 - 1$ counts the number of generators.) It is true in general that $T_{\bar{N}}^I = -(T_N^I)^*$. Since the generators are hermitian, we have $(T_N^I)^* = (T_N^I)^T$, or in components $[(T_N^I)^a{}_b]^* = (T_N^I)_b{}^a$. Therefore:

$$(T_{\bar{N}}^I)^{\bar{a}}{}_{\bar{b}}\psi_a{}^{\bar{b}} = -[(T_N^I)^{\bar{a}}{}_{\bar{b}}]^*\psi_a{}^{\bar{b}} = -(T_N^I)_{\bar{b}}{}^{\bar{a}}\psi_a{}^{\bar{b}} = -\psi_a{}^{\bar{b}}(T_N^I)_{\bar{b}}{}^{\bar{a}}$$

where the last equality is simply to bring the indices in matrix multiplication order. The covariant derivative is therefore

$$(\mathcal{D}_\mu \psi)_a{}^{\bar{a}} = \partial_\mu \psi_a{}^{\bar{a}} + i \sum_{I=1}^{N^2-1} A_\mu^I \left[(T_N^I)_a{}^b \psi_b{}^{\bar{a}} - \psi_a{}^{\bar{b}}(T_N^I)_{\bar{b}}{}^{\bar{a}}\right]\ .$$

Suppressing the matrix indices in the usual way and defining the matrix-valued gauge field $A_\mu \equiv \sum_I A_\mu^I T_N^I$, we have

$$\mathcal{D}_\mu \psi = \partial_\mu \psi + iA_\mu \psi - i\psi A_\mu$$

208

which matches our lattice covariant derivative $\mathcal{D}U = \partial U + iAU - iUA'$.

Now we return to the problem. Since $(\mathcal{D}U)^\dagger = \partial U^\dagger - iU^\dagger A + iA'U^\dagger$, we have

$$
\begin{aligned}
&\mathrm{tr}(\mathcal{D}U)(\mathcal{D}U)^\dagger \\
&\quad = \mathrm{tr}\left[\partial U \partial U^\dagger - i\partial U(U^\dagger A - A'U^\dagger) + i(AU - UA')\partial U^\dagger + (AU - UA')(U^\dagger A - A'U^\dagger)\right] \\
&\quad = \mathrm{tr}\left[\partial U \partial U^\dagger\right] - i\,\mathrm{tr}\left[(\partial U U^\dagger - U\partial U^\dagger)A + (\partial U^\dagger U - U^\dagger \partial U)A'\right] \\
&\qquad + \mathrm{tr}\left[A^2 - 2AUA'U^\dagger + A'^2\right]
\end{aligned}
$$

Recall our previous result $\mathrm{tr}\,U_{ij}U_{jk}U_{k\ell}U_{\ell i} = \text{const} +$

$$
-\frac{1}{2}a^2\mathrm{tr}\left\{\partial U \partial U^\dagger - i\left[(\partial U^\dagger U - U^\dagger\partial U)A' + (\partial U U^\dagger - U\partial U^\dagger)A\right] + A'^2 - 2AUA'U^\dagger + A^2\right\} .
$$

Therefore $\mathrm{tr}\,U_{ij}U_{jk}U_{k\ell}U_{\ell i} = -\frac{1}{2}a^2\mathrm{tr}[(\mathcal{D}U)(\mathcal{D}U)^\dagger]$ for $ijk\ell$ bounding a rectangular plaquette. The full lattice gauge theory action

$$
\mathcal{S} = \sum_{P\,\epsilon\,\text{square}} \mathrm{Re\,tr}\,U_{ij}U_{jk}U_{k\ell}U_{\ell i} + \frac{a^2}{b^2}\sum_{P\,\epsilon\,\text{rectangle}} \mathrm{Re\,tr}\,U_{ij}U_{jk}U_{k\ell}U_{\ell i}
$$

in the continuum limit $a \to 0$ with b held fixed becomes

$$
\mathcal{S} = a^4 \sum_{i=1}^{\mathcal{N}} \int d^D x\,\mathrm{tr}\left[-\frac{1}{4}(F_i)_{\mu\nu}(F_i)^{\mu\nu} - \frac{1}{2}(\mathcal{D}_\mu U)_i(\mathcal{D}^\mu U)_i^\dagger\right]
$$

where $b\mathcal{N}$ is the length of the lattice in the extra dimension. This is the action of $\mathcal{N}$ copies of a D-dimensional continuous $SU(N)$ Yang-Mills theory with a Lorentz-scalar U transforming under the $(N \otimes \bar{N})$-dimensional representation. For further discussion in the context of Kaluza-Klein theory, see http://arxiv.org/pdf/hep-th/0104005v1.

4. Study the alternative limit $b \to 0$ with a kept fixed so that you obtain a theory on a spatial lattice but with continuous time.

The result of the previous problem carries over immediately:

$$
\mathcal{S} = b^4 \sum_{i=1}^{\mathcal{N}} \int dt\,\mathrm{tr}\left[-\frac{1}{4}(F_i)_{\mu\nu}(F_i)^{\mu\nu} - \frac{1}{2}(\mathcal{D}_\mu U)_i(\mathcal{D}^\mu U)_i^\dagger\right]
$$

where $a^D\mathcal{N}$ is the volume of the lattice. We now have $\mathcal{N}$ copies of a 1-dimensional $SU(N)$ gauge theory with a Lorentz-scalar $U \sim N \otimes \bar{N}$.

5. Show that for lattice gauge theory the Wilson area law holds in the limit of strong coupling. [Hint: Expand (20) in powers of f^{-2}.]
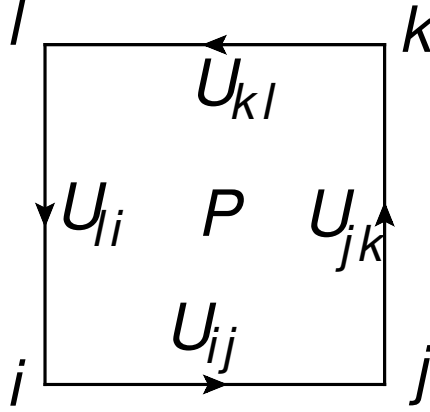
$$
\langle W(C)\rangle = \frac{1}{\mathcal{Z}}\int \left[\prod dU\right] e^{-\frac{1}{2f^2}\sum_P S(P)} W(C) \qquad (20)
$$

Solution:

Consider an $SU(N)$ lattice gauge theory whose partition function is[19]

$$\mathcal{Z} = \int \mathcal{D}U \, e^{\beta \sum_P S_P} \,, \quad S_P = \frac{1}{2N} \operatorname{tr}\left(K_P + K_P^\dagger\right)$$

where $\beta \equiv 2N/g^2$, and the function K_P is defined as $K_P \equiv U_{ij}U_{jk}U_{k\ell}U_{\ell i}$ for a plaquette P whose four corners are the lattice sites x_i, x_j, x_k, x_ℓ as in the diagram below:

Now consider a large closed rectangular curve C on the lattice whose sides have lengths R and T. Define the Wilson loop W_C for the curve C as

$$W_C = \operatorname{tr} U_{ij}U_{jk}...U_{mn}U_{ni}$$

where $x_i \to x_j \to x_k \to ... \to x_m \to x_n \to x_i$ are the links that comprise the curve C. We are interested in the expectation value of the Wilson loop:

$$\langle W_C \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}U \, W_C \, e^{\beta \sum_P S_P}$$

$$= \frac{1}{\mathcal{Z}} \int \mathcal{D}U \, W_C \prod_P \left[\sum_{n=0}^{\infty} \frac{1}{n!} (\beta S_P)^n \right]$$

---

[19]We have written the theory in this form to suggest a connection to statistical mechanics with temperature $T \equiv \beta^{-1}$. Our "strong-coupling expansion" is analogous to the high-temperature expansion $T \to \infty$ or $\beta \to 0$.

To find the leading contribution in the limit of small β, we have to understand how to carry out integrals over the link variables U_{ij}. The rules are as follows[20]:

$$\int \mathcal{D}U \, U_A{}^B = 0$$

$$\int \mathcal{D}U \, U_A{}^B \, U_C{}^D = 0$$

$$\int \mathcal{D}U \, U_A{}^B \, (U^\dagger)_C{}^D = \frac{1}{N} \delta_A^D \delta_C^B$$

$$\int \mathcal{D}U \, U_{A_1}^{B_1} \, U_{A_2}^{B_2} \, ... \, U_{A_N}^{B_N} = \frac{1}{N!} \varepsilon_{A_1 A_2 ... A_N} \varepsilon^{B_1 B_2 ... B_N}$$

for any N-by-N link variable U. (The indices $A, B, ...$ run from 1 to N.) To get the leading nonzero contribution to $\langle W_C \rangle$, we need to pair up each U_{ij} in W_C with exactly one $U_{ij}^\dagger$ from $\prod_{(P \text{ in } \Sigma)} \sum_n \frac{1}{n!} (\beta S_P)^n$, where Σ is the surface of minimal area whose boundary is the curve C. Thus the leading contribution from the product over plaquettes in Σ is when $n = 1$ in the sum above, and the integral can be approximated to leading order as[21]

$$\langle W_C \rangle \approx \int \mathcal{D}U \, W_C \, \left(\frac{\beta}{2N} \right)^{n_*} (\mathrm{tr} K_1^\dagger) ... (\mathrm{tr} K_{n_*}^\dagger)$$

where $n_* \equiv RT/a^2$ is the number of plaquettes that fit inside the region Σ, and a^2 is the area of a square plaquette with sides of length a:
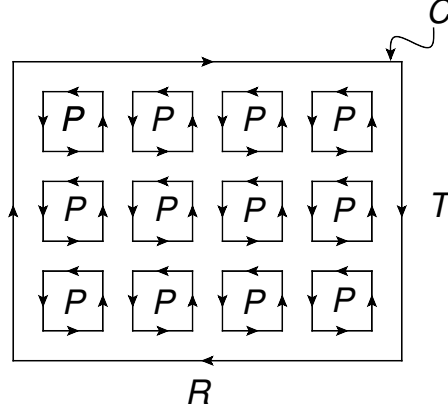
To understand how to compute the integral, first consider the overly simplified but illustrative case for which the curve C encloses only 1 plaquette, bounded by the lattice sites (i, j, k, ℓ) as in the diagram above. In this case, the Wilson loop is simply

$$W_C = \mathrm{tr} \, U_{ij} U_{jk} U_{k\ell} U_{\ell i} = (U_{ij})_A{}^B (U_{jk})_B{}^C (U_{k\ell})_C{}^D (U_{\ell i})_D{}^A$$

---

[20]See p. 90 and p. 222 of the text "Lattice Gauge Theories - An Introduction, 2nd Ed." by Heinz. J. Rothe for more details.

[21]The denominator $\mathcal{Z}$ contributes only a factor of 1 to this order. To the next order in β, we have to worry about nontrivial contributions from $\mathcal{Z}$.

and the only K that contributes is $K_1 = \operatorname{tr} U_{ij} U_{jk} U_{k\ell} U_{\ell i}$, so that

$$K_1^\dagger = U_{\ell i}^\dagger U_{k\ell}^\dagger U_{jk}^\dagger U_{ij}^\dagger = (U_{\ell i}^\dagger)_E^{\ F} (U_{k\ell}^\dagger)_F^{\ G} (U_{jk}^\dagger)_G^{\ H} (U_{ij}^\dagger)_H^{\ E} \ .$$

The expectation value of the Wilson loop to leading order[22] is

$$\langle W_C \rangle \approx \frac{\beta}{2N} \int \mathcal{D}U \, (U_{ij})_A^{\ B} (U_{jk})_B^{\ C} (U_{k\ell})_C^{\ D} (U_{\ell i})_D^{\ A} \ (U_{\ell i}^\dagger)_E^{\ F} (U_{k\ell}^\dagger)_F^{\ G} (U_{jk}^\dagger)_G^{\ H} (U_{ij}^\dagger)_H^{\ E}$$

$$= \frac{\beta}{2N} \left[\int \mathcal{D}U_{ij} (U_{ij})_A^{\ B} (U_{ij}^\dagger)_H^{\ E} \right] \left[\int \mathcal{D}U_{jk} (U_{jk})_B^{\ C} (U_{jk}^\dagger)_G^{\ H} \right] \times$$

$$\left[\int \mathcal{D}U_{k\ell} (U_{k\ell})_C^{\ D} (U_{k\ell}^\dagger)_F^{\ G} \right] \times \left[\int \mathcal{D}U_{\ell i} (U_{\ell i})_D^{\ A} (U_{\ell i}^\dagger)_E^{\ F} \right]$$

$$= \frac{\beta}{2N} \left[\frac{1}{N} \delta_A^E \delta_H^B \right] \left[\frac{1}{N} \delta_B^H \delta_G^C \right] \left[\frac{1}{N} \delta_C^G \delta_F^D \right] \left[\frac{1}{N} \delta_D^F \delta_E^A \right]$$

$$\frac{\beta}{2N} \left[\frac{1}{N} \delta_A^A \frac{1}{N} \delta_B^B \frac{1}{N} \delta_C^C \frac{1}{N} \delta_D^D \right]$$

$$= \frac{\beta}{2N} \ .$$

If we now consider the second simplest class of example, for which the curve C contains exactly two plaquettes, we will find $\langle W_C \rangle \approx \left(\frac{\beta}{2N} \right)^2$, where again all of the factors of $\frac{1}{N}$ will cancel. As soon as the curve C is large enough such that there can actually be an interior to the surface Σ – that is, for which some plaquettes will not be along the perimeter C – then the cancellation of $\frac{1}{N}$ factors will be incomplete. After the integrations, there will be $n_* - 1 = RT/a^2 - 1$ factors of $(1/N)$ left over, so the leading contribution to the expectation value of the Wilson loop is

$$\langle W_C \rangle \approx \left(\frac{\beta}{2N} \right)^{RT/a^2} \left(\frac{1}{N} \right)^{RT/a^2 - 1} = N \left(\frac{\beta}{2N^2} \right)^{RT/a^2} \ .$$

Recalling the definition $\beta = 2N/g^2$, we obtain

$$\langle W_C \rangle \approx N \left(\frac{1}{g^2 N} \right)^{RT/a^2} = N \, e^{-\frac{RT}{a^2} \ln(g^2 N)}$$

The behavior $\langle W_C \rangle \sim e^{-(\#)RT}$ is called the area law. As explained on p. 377, the energy between a quark and an antiquark is

$$E(R) = -\frac{1}{T} \ln \langle W_C \rangle = \sigma \, R + O(1/T)$$

where we have defined the string tension

$$\sigma = \frac{\ln(g^2 N)}{a^2} \ .$$

We can neglect the constant of order $1/T$ in the limit of taking the curve C arbitrarily large.

---

[22]Remember that the integral $\int \mathcal{D}U$ is written in a condensed notation to denote integration over *all* link variables. So $\int \mathcal{D}U = \int [\prod_{\langle ij \rangle} \mathcal{D}U_{ij}]$, where U_{ij} connects the sites i and j on the lattice and $\langle ij \rangle$ denotes that sites i and j are nearest neighbors.

VII.2   Electroweak Unification

1. Unfortunately, the mass of the elusive Higgs particle H depends on the parameters in the double well potential $V = -\mu^2\varphi^\dagger\varphi + \lambda(\varphi^\dagger\varphi)^2$ responsible for the spontaneous symmetry breaking. Assuming that H is massive enough to decay into W^+W^- and ZZ, determine the rates for H to decay into various modes.

Solution:

We compute the rates for $H \to W^+W^-$, $H \to ZZ$ and $H \to \ell^+\ell^-$ at tree level. We fix unitary gauge, which is the statement that we write the Higgs doublet φ as

$$\varphi = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + H) \end{pmatrix}$$

where H is a real scalar field whose quantum is the spin-0 Higgs boson. As described on p. 382, after using the above parameterization for φ in the Lagrangian $\mathcal{L} = (D_\mu\varphi)^\dagger(D^\mu\varphi)$ and identifying the photon A and the Z-boson through the rotation

$$\begin{pmatrix} Z \\ A \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} W^3 \\ B \end{pmatrix}$$

you should eventually discover an interaction Lagrangian

$$\mathcal{L} = -\frac{2}{v}H(M_W^2 W^{+\mu}W_\mu^- + \frac{1}{2}M_Z^2 Z^\mu Z_\mu)$$

Including the field H in addition to the constant v in equation (2) on p. 381 yields the Higgs-lepton interaction

$$\mathcal{L} = -\frac{m_\ell}{v}H\bar{\ell}_L\ell_R + h.c. = -\frac{m_\ell}{v}H\bar{\ell}\ell$$

where ℓ is the 4-component Dirac field for either an electron, a muon or a tau. These three interactions yield the following vertices:

$$(HWW) = -2iv^{-1}M_W^2 g_{\mu\nu} \quad, \quad (HZZ) = -2iv^{-1}M_Z^2 g_{\mu\nu} \quad, \quad (H\ell\ell) = -im_\ell v^{-1}$$

Note the factor of 2 in the HZZ vertex due to the interchange symmetry of the Zs.

The amplitude $\mathcal{M}$ for $H \to ZZ$ is

$$\mathcal{M} = -2iv^{-1}M_Z^2 g_{\mu\nu}\varepsilon_{\lambda_1}^\mu(k_1)\varepsilon_{\lambda_2}^\nu(k_2)$$

which implies

$$|\mathcal{M}|^2 = 4\left(\frac{M_z^2}{v}\right)^2 \left[\varepsilon_{\lambda_1}^\mu(k_1)\varepsilon_{\lambda_1}^\nu(k_1)^*\right]\left[\varepsilon_{\lambda_2\mu}(k_2)\varepsilon_{\lambda_2\nu}(k_2)^*\right]$$

Summing over outgoing polarizations using the rule

$$\sum_\lambda \varepsilon_\lambda^\mu(k)\varepsilon_\lambda^\nu(k)^* = g^{\mu\nu} - \frac{k^\mu k^\nu}{M_Z^2}$$

gives the polarization-averaged amplitude-squared $\mathbb{M}^2 \equiv \sum_{\lambda_1,\lambda_2}|\mathcal{M}|^2$:

$$\begin{aligned}
\mathbb{M}^2 &= 4\left(\frac{M_z^2}{v}\right)^2\left(g^{\mu\nu} - \frac{k_1^\mu k_1^\nu}{M_Z^2}\right)\left(g_{\mu\nu} - \frac{k_{2\mu}k_{2\nu}}{M_Z^2}\right)\\
&= 4\left(\frac{M_z^2}{v}\right)^2\left[4 - \frac{1}{M_Z^2}(k_1^2 + k_2^2) + \left(\frac{k_1^\mu k_{2\mu}}{M_Z^2}\right)^2\right]\\
&= 4\left(\frac{M_z^2}{v}\right)^2\left[2 + \left(\frac{k_1^0 k_2^0 - \vec{k}_1\cdot\vec{k}_2}{M_Z^2}\right)^2\right]
\end{aligned}$$

where in the last line we have used the fact that the outgoing Z bosons are on-shell, meaning $k_1^2 = k_2^2 = M_Z^2$. Now we use equation (38) on p. 141 for the differential decay rate in the center of mass frame:

$$d\Gamma = \frac{1}{2m_H}\frac{d^3k_1}{(2\pi)^3 2k_1^0}\frac{d^3k_2}{(2\pi)^3 2k_2^0}(2\pi)^4\delta^{(4)}(K - k_1 - k_2)\,\mathbb{M}^2$$

where $m_H \equiv \sqrt{\lambda}\,v$ is the mass of the Higgs boson and K is its 4-momentum. When integrating over the outgoing momenta, the fact that there are two identical outgoing Z bosons forces us to divide by 2, meaning that the total decay rate is $\Gamma = \frac{1}{2}\int d\Gamma$. Now we compute:

$$\begin{aligned}
\Gamma &= \frac{1}{2}\left(\frac{1}{2m_H}\right)\left(\frac{1}{2^2(2\pi)^2}\right)\int\frac{d^3k_1}{k_1^0}\int\frac{d^3k_2}{k_2^0}\delta(m_H - k_1^0 - k_2^0)\delta^{(3)}(\vec{k}_1 + \vec{k}_2)\,\mathbb{M}^2\\
&= \frac{1}{2^6\pi^2 m_H}\int\frac{4\pi\,dk\,k^2}{k^2 + M_Z^2}\delta(m_H - 2\sqrt{k^2 + M_Z^2})\,\mathbb{M}^2
\end{aligned}$$

Define $f(k) \equiv m_H - 2\sqrt{k^2 + M_Z^2}$ and use the formula $\delta(k) = \frac{1}{|f'(k_*)|}\delta(k - k_*)$, where k_* is defined by $f(k_*) \equiv 0$. $f(k) = 0 \implies k^2 = \frac{m_H^2}{4} - M_Z^2$. The derivative of $f(k)$ at this value of k is

$$f'(k) = \frac{-2k}{\sqrt{k^2 + M_Z^2}} = -\frac{4k}{m_H}$$

where we wait until later to plug in the value for k. Using this, we have

$$\begin{aligned}
\Gamma &= \frac{1}{2^4\pi m_H}\frac{k^2}{k^2 + M_Z^2}\frac{m_H}{4k}\mathbb{M}^2\\
&= \frac{1}{2^4\pi}\frac{k}{4(k^2 + M_z^2)}\mathbb{M}^2\\
&= \frac{1}{2^4\pi}\frac{\sqrt{m_H^2/4 - M_Z^2}}{m_H^2}\mathbb{M}^2\\
&= \frac{1}{2^5\pi m_H}\sqrt{1 - \left(\frac{2M_Z}{m_H}\right)^2}\,\mathbb{M}^2
\end{aligned}$$

214

If the Z bosons were spin-0 particles, then there would be no momentum dependence in $\mathbb{M}$ and we would get the behavior $\Gamma \propto \sqrt{1 - (2M_Z/m_H)^2}$ we expect from the discussion of $1 \to 2$ meson decay on p. 142. The decay rate is only real if $m_H > 2M_Z$, which reflects the perfectly sensible fact that the Higgs can only decay into two Z bosons if its mass is larger than twice the mass of the Z boson.

Now let's see what the spin-1 polarizations do. Using the integrations over the delta functions, we have:

$$(k_1^0 k_2^0 - \vec{k}_1 \cdot \vec{k}_2)^2 = [(k^2 + M_Z^2) + k^2]^2 = \left[\frac{m_H^2}{4} + \frac{m_H^2}{4} - M_Z^2 \right]^2 = \frac{m_H^4}{4} \left(1 - \frac{2M_Z^2}{m_H^2} \right)^2$$

Therefore:

$$\Gamma = \frac{1}{2^5 \pi m_H} \sqrt{1 - \left(\frac{2M_Z}{m_H} \right)^2} \, 4 \frac{M_Z^4}{v^2} \left[2 + \frac{m_H^4}{4M_Z^4} \left(1 - \frac{2M_Z^2}{m_H^2} \right)^2 \right]$$

$$= \frac{m_H}{4\pi} \left(\frac{M_Z^2}{m_H v} \right)^2 \sqrt{1 - \left(\frac{2M_Z}{m_H} \right)^2} \left[1 + \frac{m_H^4}{8M_Z^4} \left(1 - \frac{2M_Z^2}{m_H^2} \right)^2 \right]$$

It is often convenient to display these decay rates as polynomials in the variable $x \equiv 4M_Z^2/m_H^2$. Substituting $M_Z^2 = xm_H^2/4$ gives

$$1 + \frac{m_H^4}{32M_Z^4} \left(1 - \frac{2M_Z^2}{m_H^2} \right)^2 = 1 + \frac{m_H^4}{8(xm_H^2/4)^2} \left(1 - \frac{2(xm_H^2/4)}{m_H^2} \right)^2$$

$$= 1 + \frac{2}{x^2} \left(1 - \frac{x}{2} \right)^2 = \frac{2}{x^2} \left[\frac{x^2}{2} + \left(1 - \frac{x}{2} \right)^2 \right]$$

$$= \frac{2}{x^2} \left(\frac{x^2}{2} + 1 - x + \frac{x^2}{4} \right) = \frac{1}{2x^2} (3x^2 - 4x + 4)$$

Therefore the decay rate is

$$\Gamma(H \to ZZ) = \frac{m_H^3}{2^7 \pi v^2} \sqrt{1 - x} \, (3x^2 - 4x + 4) \quad , \quad x \equiv \frac{4M_Z^2}{m_H^2}$$

When comparing to other sources[23] substitute v for the Fermi constant $G_F \equiv 1/(v^2\sqrt{2})$, although be aware that different definitions of G_F exist, for example without the $\sqrt{2}$. To get an idea for the size of the prefactor, take $m_H = v \approx 246$ GeV, for which $m_H^3/(2^7 \pi v^2) \sim 0.6$ GeV and $x \sim 0.5$ so that $\Gamma \sim 1.8$ GeV.

To calculate the decay rate for $H \to W^+ W^-$, note that almost everything would go through in exactly the same way as for $H \to ZZ$ with M_Z replaced by M_W. The only difference is

---

[23]For example, B. W. Lee, C. Quigg and H. B. Thacker, "Weak Interactions at Very High Energies: the Role of the Higgs Boson Mass," FERMILAB-Pub-77/30-THY March 1977.

that since W^+ and W^- are not identical particles, there is no factor of $1/2$ when integrating $d\Gamma$ to get Γ. Therefore we can immediately write down

$$\Gamma(H \to W^+W^-) = \frac{m_H^3}{2^6\pi v^2}\sqrt{1-y}\,(3y^2 - 4y + 4) \ , \quad y \equiv \frac{4M_W^2}{m_H^2}$$

Note that $\Gamma(H \to W^+W^-)/\Gamma(H \to ZZ)$ is not quite 2, since $M_W/M_Z < 1$.

Finally, we compute the decay width of the Higgs boson into a lepton pair $\ell^+\ell^-$, which becomes the dominant decay channel if $m_H < 2M_W \sim 190$ GeV. The vertex is

$$(H\bar{\ell}\ell) = imv^{-1}$$

so the amplitude is simply $\mathcal{M} = imv^{-1}\bar{u}(p_1,s_1)v(p_2,s_2)$, which implies that the spin-summed amplitude squared $\mathbb{M}^2 \equiv \sum_{s_1,s_2}|\mathcal{M}|^2$ is

$$\mathbb{M}^2 = \frac{m^2}{v^2}\text{tr}\left[\left(\frac{\not{p}_1 + m}{2m}\right)\left(\frac{\not{p}_2 - m}{2m}\right)\right] = \frac{1}{v^2}(\,p_1^\mu\, p_{2\mu} - m^2)$$

The integrals over the momentum-conserving delta function will work out in exactly the same way as for the $H \to ZZ$ mode, which implies $p_1^\mu p_{2\mu} - m^2 = 2p^2$, where $p \equiv |\vec{p}_1| = |\vec{p}_2|$. Therefore, as far as the decay rate is concerned, we have

$$\mathbb{M}^2 = \frac{4p^2}{v^2}$$

The decay rate formula is

$$\Gamma = \frac{1}{2m_H}\int\frac{d^3p_1}{(2\pi)^3(p_1^0/m)}\int\frac{d^3p_2}{(2\pi)^3(p_2^0/m)}(2\pi)^4\delta(m_H - p_1^0 - p_2^0)\delta^{(3)}(\vec{p}_1 + \vec{p}_2)\,\mathbb{M}^2$$

We write the formula explicitly to remind you of the factor p^0/m for the fermions, although everything else will proceed in the same way as before, leaving the result

$$\Gamma(H \to \ell^+\ell^-) = \frac{m_\ell^2}{16\pi v^2}m_H\left(1 - \frac{4m_\ell^2}{m_H^2}\right)^{3/2}$$

Note that $m_\ell = fv/\sqrt{2}$, where f is defined by the Lagrangian $\mathcal{L} = f\,\bar{\psi}_L\varphi\,\ell_R + h.c.$ (using the notation of pp. 380-381), so the limit $m_\ell \to 0$ is perfectly fine and just leaves behind

$$\Gamma(H \to \ell^+\ell^-) \to \frac{f^2}{32\pi}m_H \ .$$

2. Show that it is possible to stay with the $SU(2)$ gauge group and to identify W^3 as the photon A, but at the cost of inventing some experimentally unobserved lepton fields. This theory does not describe our world: For one thing, it is essentially impossible to incorporate the quarks. Show this! [Hint: We have to put the leptons into a triplet of $SU(2)$

instead of a doublet.]

Solution:

We take the gauge group to be $G = SU(2)$, and we will use 2-component spinor notation. Consider a lepton triplet written as a 2-by-2 symmetric matrix:

$$\ell \sim 2 \otimes_S 2 \implies \ell_{ij} = \begin{pmatrix} \ell^+ & \frac{1}{\sqrt{2}}\ell^0 \\ \frac{1}{\sqrt{2}}\ell^0 & \ell^- \end{pmatrix}.$$

The superscript labels indicate that we are anticipating a bit and denoting the electric charge of each component. With the gauge bosons $(W_\mu)_i{}^j \equiv \sum_{a=1}^3 W_\mu^a(\frac{1}{2}\sigma^a)_i{}^j$, the covariant derivative acting on ℓ_{ij} with gauge coupling g is

$$i(D_\mu\ell)_{ij} = i\partial_\mu\ell_{ij} + g\begin{pmatrix} W_\mu^3\ell^+ + W_\mu^+\ell^0 & \frac{1}{\sqrt{2}}(W_\mu^+\ell^- + W_\mu^-\ell^+) \\ \frac{1}{\sqrt{2}}(W_\mu^+\ell^- + W_\mu^-\ell^+) & -W_\mu^3\ell^- + W_\mu^-\ell^0 \end{pmatrix}.$$

The currents defined by $\mathcal{L} = i\ell^\dagger\bar\sigma^\mu D_\mu\ell = i\ell^\dagger\bar\sigma^\mu\partial_\mu\ell + W_\mu^3 J_3^\mu + W_\mu^+ J^{-\mu} + W_\mu^- J^{+\mu}$ are therefore:

$$J_3^\mu = g[(\ell^+)^\dagger\bar\sigma^\mu\ell^+ - (\ell^-)^\dagger\bar\sigma^\mu\ell^-]$$
$$J^{-\mu} = g[(\ell^+)^\dagger\bar\sigma^\mu\ell^0 + (\ell^0)^\dagger\bar\sigma^\mu\ell^-]$$
$$J^{+\mu} = g[(\ell^0)^\dagger\bar\sigma^\mu\ell^+ + (\ell^-)^\dagger\bar\sigma^\mu\ell^0]$$

If W_μ^3 remains massless, then we see that J_3^μ is the correct definition of the electromagnetic current, with the electric coupling defined as $e \equiv g$.

Consider a Higgs field $\phi \sim 2 \otimes_S 2$. Its covariant derivative is the same as that for ℓ, and its interactions with the gauge bosons comes from the kinetic term $\mathcal{L} = (D^\mu\phi)^\dagger D_\mu\phi$. When ϕ^0 obtains a vacuum expectation value, we find

$$(D_\mu\phi)_{ij} = -ig\langle\phi^0\rangle\begin{pmatrix} W_\mu^+ & 0 \\ 0 & W_\mu^- \end{pmatrix} + \dots$$

and therefore

$$(D^\mu\phi)^\dagger D_\mu\phi = 2g^2|\langle\phi^0\rangle|^2 W_\mu^+ W^{-\mu} + \dots$$

The charged bosons $W^\pm$ get a mass $m_W = \sqrt{2}g|\langle\phi^0\rangle|$ and the neutral boson W^3 remains massless. Given this and the lepton currents derived previously, we can indeed identify W^3 as the photon.

However, there is a problem in assigning masses to the charged leptons while keeping the neutral one massless (or at least approximately massless). From $\ell_{ij} \sim 2 \otimes_S 2$, we can form the singlet

$$\tfrac{1}{2}\ell_{ij}\varepsilon^{ik}\varepsilon^{j\ell}\ell_{k\ell} = -\tfrac{1}{2}\ell^0\ell^0 + \ell^-\ell^+ \ .$$

If we couple an $SU(2)$-singlet Higgs φ with a nonvanishing vacuum expectation value $\langle\varphi\rangle$ to this term, then we find a Majorana mass for the neutral lepton with the same magnitude as the Dirac mass for the charged lepton. This is clearly incompatible with the near masslessness of the neutrino that participates in nuclear beta decay.

For further discussion as well as for other models, we refer the reader to the literature.

Addendum: Covariant Derivative for the Triplet

Here we derive the electroweak $SU(2) \otimes U(1)$ covariant derivatives for a field transforming as $\phi \sim (2 \otimes_S 2 = 3, y)$, for arbitrary values of the hypercharge y. (Here we will use the convention for which the hypercharge generator is denoted by Y rather than $\tfrac{1}{2}Y$.)

For the generators of $SU(2)$, we have:

$$W^a\left(\frac{\sigma^a}{2}\right) = \frac{1}{2}\begin{pmatrix} W^3 & \sqrt{2}W^+ \\ \sqrt{2}W^- & -W^3 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} cZ + sA & \sqrt{2}W^+ \\ \sqrt{2}W^- & -(cZ + sA) \end{pmatrix}$$

As usual, we define the neutral boson Z and the photon A as a linear combination of W^3 and B:

$$\begin{pmatrix} Z \\ A \end{pmatrix} \equiv \begin{pmatrix} c & -s \\ s & c \end{pmatrix}\begin{pmatrix} W^3 \\ B \end{pmatrix} \implies \begin{pmatrix} W^3 \\ B \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} Z \\ A \end{pmatrix} = \begin{pmatrix} cZ + sA \\ -sZ + cA \end{pmatrix}$$

where $s \equiv \sin\theta$ and $c \equiv \cos\theta$ define the weak mixing angle $s/c \equiv g'/g$, where g is the $SU(2)$ gauge coupling and g' is the $U(1)$ gauge coupling. This implies

$$\frac{s}{c}B = s(-\frac{s}{c}Z + A)$$

which is a combination that appears in the covariant derivative. We have:

$$(D\phi)_{ij} = \partial\phi_{ij} - i\frac{g}{2}[W^a(\sigma^a)_i{}^{i'}\delta_j{}^{j'} + W^a\delta_i{}^{i'}(\sigma^a)_j{}^{j'} + 2y\frac{s}{c}B\delta_i{}^{i'}\delta_j{}^{j'}]\phi_{i'j'} \ .$$

In detail:

$$\begin{aligned} i[(D - \partial)\phi]_{11} &= +\frac{g}{2}[W^a(\sigma)_1{}^i\delta_1{}^j + W^a\delta_1{}^i(\sigma^a)_1{}^j + 2y\frac{s}{c}B\delta_1{}^i\delta_1{}^j]\phi_{ij} \\ &= \frac{g}{2}[W^a(\sigma^a)_1{}^i\phi_{i1} + W^a(\sigma^a)_1{}^j\phi_{1j} + 2ys(-\frac{s}{c}Z + A)\phi_{11}] \\ &= \frac{g}{2}[2(cZ + sA)\phi_{11} + \sqrt{2}W^+(\phi_{12} + \phi_{21}) + 2ys(-\frac{s}{c}Z + A)\phi_{11}] \\ &= g\left\{[c(1 - y(\frac{s}{c})^2)Z + s(1 + y)A]\phi_{11} + \sqrt{2}W^+\phi_{12}\right\} \end{aligned}$$

$$i[(D - \partial)\phi]_{12} = \frac{g}{2}[W^a(\sigma^a)_1{}^i\delta_2{}^j + W^a\delta_1{}^i(\sigma^a)_2{}^j + 2y\frac{s}{c}B\delta_1^i\delta_2^j]\phi_{ij}$$
$$= \frac{g}{2}[W^a(\sigma^a)_1{}^i\phi_{i2} + W^a(\sigma^a)_2{}^j\phi_{1j} + 2y\frac{s}{c}B\phi_{12}]$$
$$= \frac{g}{2}[(W^3 - W^3)\phi_{12} + \sqrt{2}W^+\phi_{22} + \sqrt{2}W^-\phi_{11} + 2ys(-\frac{s}{c}Z + A)\phi_{12}]$$
$$= g[ys(-\frac{s}{c}Z + A)\phi_{12} + \frac{1}{\sqrt{2}}(W^+\phi_{22} + W^-\phi_{11})]$$

$$i[(D - \partial)\phi]_{22} = \frac{g}{2}[W^a(\sigma^a)_2{}^1\phi_{12} + W^a(\sigma^a)_2{}^2\phi_{22} + W^a(\sigma^a)_2{}^1\phi_{21} + W^a(\sigma^a)_2{}^2\phi_{22} + 2y\frac{s}{c}B\phi_{22}]$$
$$= g\left\{\sqrt{2}W^-\phi_{12} + [-c(1 + y(\frac{s}{c})^2)Z + s(-1 + y)A]\right\}$$

To summarize:

$$[i(D - \partial)\phi]_{ij} =$$
$$g\begin{pmatrix} [c(1 - y\frac{s^2}{c^2})Z + s(1 + y)A]\phi_{11} + \sqrt{2}W^+\phi_{12} & ys(-\frac{s}{c}Z + A)\phi_{12} + \frac{1}{\sqrt{2}}(W^+\phi_{22} + W^-\phi_{11}) \\ \times & [-c(1 + y\frac{s^2}{c^2})Z + s(-1 + y)A]\phi_{22} + \sqrt{2}W^-\phi_{12} \end{pmatrix}$$

Let us check that the couplings to the photon come out as planned. The generator of electric charge is $Q = T^3 + Y$, or in components $Q_{ij}{}^{i'j'} = \left(\frac{\sigma^3}{2}\right)_i{}^{i'}\delta_j{}^{j'} + \delta_i{}^{i'}\left(\frac{\sigma^3}{2}\right)_j{}^{j'} + y\delta_i^{i'}\delta_j^{j'}$. We find:

$$(Q\phi)_{11} = [(+\tfrac{1}{2}) + (+\tfrac{1}{2}) + y]\phi_{11} = (1 + y)\phi_{11} \ \checkmark$$
$$(Q\phi)_{12} = [(+\tfrac{1}{2}) + (-\tfrac{1}{2}) + y]\phi_{12} = y\phi_{12} \ \checkmark$$
$$(Q\phi)_{22} = [(-\tfrac{1}{2}) + (-\tfrac{1}{2}) + y]\phi_{22} = (-1 + y)\phi_{22} \ \checkmark$$

To fix the normalization of the field, notice that $\phi^{\dagger ij}\phi_{ij} = |\phi_{11}|^2 + 2|\phi_{12}|^2 + |\phi_{22}|^2$. So in general, for $\phi \sim (2 \otimes_S 2, y)$ of $SU(2) \otimes U(1)$, the components of ϕ_{ij} can be identified as follows:

$$\phi_{ij} = \begin{pmatrix} \phi^{(1+y)} & \frac{1}{\sqrt{2}}\phi^{(y)} \\ \frac{1}{\sqrt{2}}\phi^{(y)} & \phi^{(-1+y)} \end{pmatrix}$$

where the superscript denotes the electric charge.

For example: $y = -1$ implies

$$\phi_{ij} = \begin{pmatrix} \phi^0 & \frac{1}{\sqrt{2}}\phi^- \\ \frac{1}{\sqrt{2}}\phi^- & \phi^{--} \end{pmatrix}$$

with covariant derivative:

$$[i(D - \partial)\phi]_{ij} =$$
$$\begin{pmatrix} \frac{1}{c}Z\phi^0 + W^+\phi^- & -s(A - \frac{s}{c}Z)\frac{1}{\sqrt{2}}\phi^- + \frac{1}{\sqrt{2}}(W^+\phi^{--} + W^-\phi^0) \\ \times & [-c(1 - \frac{s^2}{c^2})Z - 2sA]\phi^{--} + W^-\phi^- \end{pmatrix}$$

This Higgs triplet can couple to the left-handed lepton triplet $\ell_{(i}\ell_{j)}$ to generate neutrino masses.

VII.3　　Quantum Chromodynamics

1. Calculate C in (9). [Hint: If you need help, consult T. Appelquist and H. Georgi, Phys. Rev. D8: 4000, 1973; and A. Zee, *Phys. Rev.* D8: 4038, 1973.]

$$(9) \qquad R(E) \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = \left(3 \sum_a Q_a^2\right)\left(1 + C\frac{2}{(11 - \frac{2}{3}n_f)\ln(E/\mu)} + \ldots\right)$$

Solution:

The details of this two-loop calculation can be found on p. 415 (chapter 8-4-4) of the quantum field theory text by Itzykson and Zuber. If you compute the integrals, you can check your answer with the Particle Data Group[24]:

$$R(E) = \left(3 \sum_a Q_a^2\right)\left[1 + \frac{\alpha_S(E)}{\pi} + O(\alpha_S^2)\right]$$

Then equation (6) on p. 389, which is

$$\alpha_S(E) = \frac{\alpha_S(\mu)}{1 + \frac{1}{2\pi}(11 - \frac{2}{3}n_f)\alpha_S(\mu)\ln(E/\mu)} \to \frac{2\pi}{(11 - \frac{2}{3}n_f)\ln(E/\mu)}$$

for large E implies $C = \pi$.

2. Calculate (2).

$$(2) \qquad \frac{dg}{dt} = -\frac{11}{3}T_2(G)\frac{g^3}{16\pi^2}$$

Solution:

We will use the background field gauge. The goal is to compute the 1-loop effective potential for pure Yang-Mills theory in the presence of a constant background gauge field.[25]

Including gauge-fixing and ghost terms (see p. 372), the full Yang-Mills Lagrangian with counterterms is

$$\mathcal{L} = \mathcal{L}_{\text{YM}} + \mathcal{L}_{\text{gf}} + \mathcal{L}_{\text{ghost}} + \mathcal{L}_{\text{ct}}$$

where

$$\mathcal{L}_{\text{YM}} = -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} , \quad F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f^{abc} A_\mu^b A_\nu^c$$

$$\mathcal{L}_{\text{gf}} = -\frac{1}{2\xi}G^a G^a$$

$$\mathcal{L}_{\text{ghost}} = (D_\mu \bar{c})^a (D^\mu c)^a , \quad (D_\mu c)^a = \partial_\mu c^a + g f^{abc} A_\mu^b c^c$$

$$\mathcal{L}_{\text{ct}} = -\frac{1}{4}(Z_g - 1)F_{\mu\nu}^a F^{a\mu\nu} + (Z_c - 1)(D_\mu \bar{c})^a (D^\mu c)^a$$

---

[24]http://pdg.ihep.su/2009/reviews/rpp2009-rev-qcd.pdf
[25]As suggested on p. 388 of the text, here we follow the calculation of S. Weinberg.

We will choose the gauge fixing function G^a shortly.

Comparing the renormalized Lagrangian $\mathcal{L} = -\frac{1}{4}Z_g F^a_{\mu\nu}F^{a\mu\nu} + ...$ to the bare Lagrangian $\mathcal{L} = -\frac{1}{4}(F_B)^a_{\mu\nu}(F_B)^{a\mu\nu} + ...$, we find the relation

$$g_B = Z_g^{1/2}g\tilde{\mu}^{\varepsilon/2}$$

between the bare coupling g_B and the renormalized coupling g. Here we have anticipated using dimensional regularization $d = 4 - \varepsilon$ to regulate the forthcoming divergent integral, and therefore replaced $g \to g\tilde{\mu}^{\varepsilon/2}$ to keep g dimensionless in $d \neq 0$. The goal is now to compute the $O(g^2)$ contribution to Z_g and thereby obtain the 1-loop beta function for the gauge coupling g.

We now split up the gauge field A^a_μ into a constant background field $\bar{A}^a_\mu$ and a fluctuation A'^a_μ: $A = \bar{A} + A'$. (See p. 504 in chapter N.3 of the text.) The full (infinitesimal) gauge transformation $\delta A^a_\mu = \partial_\mu\varepsilon^a - f^{abc}\varepsilon^b A^c_\mu$ is split up into

$$\delta\bar{A}^a_\mu = \partial_\mu\varepsilon^a - f^{abc}\varepsilon^b\bar{A}^c_\mu$$
$$\delta A'^a_\mu = -f^{abc}\varepsilon^b A'^c_\mu$$

so that $\bar{A}$ transforms as a gauge field while A' transforms as a matter field in the adjoint representation.

Putting $A = \bar{A} + A'$ into the Lagrangian and choosing the gauge-fixing function[26]

$$G^a = (\bar{D}_\mu A'^\mu)^a \equiv (\delta^{ac}\partial_\mu + gf^{abc}\bar{A}^b_\mu)A'^{c\mu}$$

we arrive at

$$\mathcal{L}_{\text{YM}} = -\frac{1}{4g^2}\left[\bar{F}^a_{\mu\nu} + (\bar{D}_\mu A'_\nu)^a - (\bar{D}_\nu A'_\mu)^a + gf^{abc}A'^b_\mu A'^c_\nu\right]^2$$
$$\mathcal{L}_{\text{gf}} = -\frac{1}{2\xi}[(\bar{D}_\mu A'^\mu)^a]^2$$
$$\mathcal{L}_{\text{ghost}} = (\bar{D}_\mu\bar{c})^a[(\bar{D}^\mu c)^a - gf^{abc}c^b A'^{c\mu}] \ .$$

Here $\bar{F}^a_{\mu\nu} = gf^{abc}\bar{A}^b_\mu\bar{A}^c_\nu$ is the field strength of the constant background field (that is, $\partial_\mu\bar{A}^a_\nu = 0$), and $\bar{D}_\mu$ is the covariant derivative with respect to the background field, as defined by the gauge fixing function G^a above.

Since $\bar{A}^a_\mu$ is to be taken as a fixed classical background field, we do not integrate over it in the path integral. The theory is invariant under the formal transformation

$$\bar{A}^a_\mu \to \bar{A}^a_\mu + \partial_\mu\varepsilon^a - gf^{abc}\varepsilon^b\bar{A}^c_\mu$$

---

[26]Note that this differs from the usual R_ξ gauge, in which the gauge fixing function is chosen as $G^a = \partial_\mu A^{a\mu}$.

irrespective of whether the transformation is performed before or after integrating over the fluctuation fields $A_\mu'^a, c^a, \bar{c}^a$. Thus by this formal background gauge invariance, we can compute the 1-loop correction to the gauge coupling by finding the coefficient of $-\frac{1}{4}\bar{F}_{\mu\nu}^a \bar{F}^{a\mu\nu} \sim (\bar{A}_\mu^a)^4$ in the 1-loop effective action, obtained by keeping only quadratic terms in $A', c, \bar{c}$ and performing the resulting gaussian integrals.

The quadratic action is

$$S_{\text{quad}} = \int d^4x \left[-\frac{1}{4}[(\bar{D}_\mu A_\nu')^a - (\bar{D}_\nu A_\mu')^a]^2 - \frac{1}{2}gf^{abc}\bar{F}^{a\mu\nu}A_\mu'^b A_\nu'^c - \frac{1}{2\xi}[(\bar{D}_\mu A'^\mu)^a]^2 + (\bar{D}_\mu \bar{c})^a(\bar{D}^\mu c)^a \right]$$

$$= \int d^4x \int d^4y \left[-\frac{1}{2}A'^{a\mu}(x)\mathcal{M}_{\mu\nu}^{ab}(x,y)A'^{b\nu}(y) + \bar{c}^a(x)\mathcal{N}^{ab}(x,y)c^b(y) \right]$$

where in the second line we have chosen the gauge $\xi = 1$ and defined the matrices

$$\mathcal{M}_{\mu\nu}^{ab}(x,y) = \eta_{\mu\nu}\left[\delta^{ca}\frac{\partial}{\partial x_\alpha} - gf^{cda}\bar{A}^{d\alpha} \right]\left[\delta^{cb}\frac{\partial}{\partial y^\alpha} - gf^{ceb}\bar{A}_\alpha^e \right]\delta^4(x-y)$$

$$- \left\{ \left[\delta^{ca}\frac{\partial}{\partial x^\nu} - gf^{cda}\bar{A}_\nu^d \right]\left[\delta^{cb}\frac{\partial}{\partial y^\mu} - gf^{ceb}\bar{A}_\mu^e \right] - (\mu \leftrightarrow \nu) \right\}\delta^4(x-y)$$

$$+ gf^{cab}\bar{F}_{\mu\nu}^c\delta^4(x-y)$$

and

$$\mathcal{N}^{ab}(x,y) = \left[\delta^{ca}\frac{\partial}{\partial x^\mu} - f^{cda}\bar{A}_\mu^d \right]\left[\delta^{cb}\frac{\partial}{\partial y_\mu} - f^{ceb}\bar{A}^{e\mu} \right]\delta^4(x-y) .$$

The one-loop effective action can now be obtained by integrating over the fluctuations A' and $c, \bar{c}$:

$$i\Gamma_{\text{1-loop}}(\bar{A}) = \int d^4x \int \frac{d^4p}{(2\pi)^4} \left[-\frac{1}{2}\text{tr}\ln M(p) + \text{tr}\ln N(p) \right]$$

where $M_{\mu\nu}^{ab}(p)$ and $N^{ab}(p)$ are defined through the Fourier transforms of $\mathcal{M}_{\mu\nu}^{ab}(x,y)$ and $\mathcal{N}^{ab}(x,y)$ respectively:

$$M_{\mu\nu}^{ab}(p) = -\eta_{\mu\nu}\left(ip_\alpha\delta^{ca} - gf^{cda}\bar{A}_\alpha^d \right)\left(ip^\alpha\delta^{cb} + gf^{ceb}\bar{A}^{e\alpha} \right)$$

$$+ \left[\left(ip_\nu\delta^{ca} - gf^{cda}\bar{A}_\nu^d \right)\left(ip_\mu\delta^{cb} + gf^{ceb}\bar{A}_\mu^e \right) - (\mu \leftrightarrow \nu) \right]$$

$$+ gf^{cab}\bar{F}_{\mu\nu}^c$$

$$N^{ab}(p) = -\left(ip_\mu\delta^{ca} - gf^{cda}\bar{A}_\mu^d \right)\left(ip^\mu\delta^{cb} + gf^{ceb}\bar{A}^{e\mu} \right) .$$

We are interested in extracting the coefficient of the term proportional to $\bar{F}_{\mu\nu}^a \bar{F}^{a\mu\nu}$ in the one-loop effective action, so we need all the terms quartic in $\bar{A}_\mu^a$. Let $\mathcal{O}_n$ denote the part of an operator $\mathcal{O}$ that is of order $(\bar{A}_\mu^a)^n$. Then the terms we are looking for are of the form:

$$(\text{tr}\ln \mathcal{O})_4 = \text{tr}\left[-\frac{1}{2}(\mathcal{O}_0^{-1}\mathcal{O}_2)^2 + (\mathcal{O}_0^{-1}\mathcal{O}_1)^2\mathcal{O}_0^{-1}\mathcal{O}_2 - \frac{1}{4}(\mathcal{O}_0^{-1}\mathcal{O}_1)^4 \right]$$

for $\mathcal{O} = M, N$. Each of these terms has a diagrammatic interpretation. The first term is
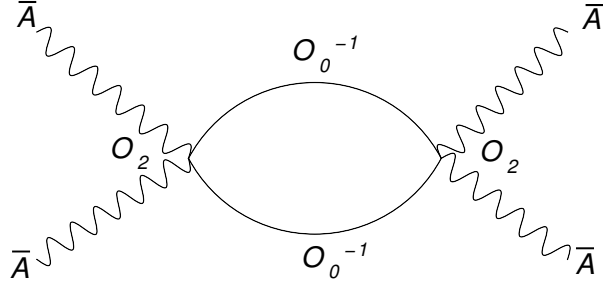
In other words, reading from right to left, the term $\mathrm{tr}(\mathcal{O}_0^{-1}\mathcal{O}_2\mathcal{O}_0^{-1}\mathcal{O}_2)$ starts with an operator with two external $\bar{A}$ lines, turns into an internal propagator $\mathcal{O}_0^{-1}$, meets another operator with two external $\bar{A}$ lines, then closes with another internal propagator $\mathcal{O}_0^{-1}$. The symmetry factors and minus signs are all automatically taken into account by the algebra at the level of the effective action.

Similarly, the second term is represented graphically as
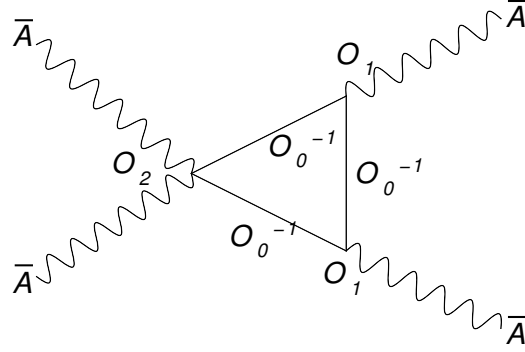
and the third term is a box diagram:
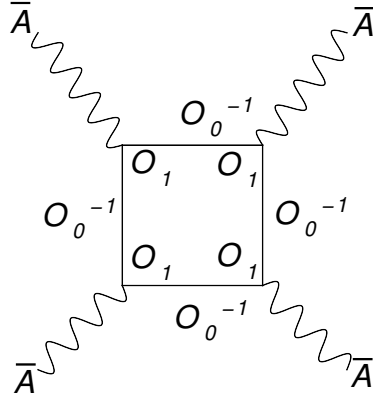
Applying this decomposition to the operators $\mathcal{O} = M$ and N and simplifying the momentum integrals using $p^\mu p^\nu \to \frac{1}{4}p^2$ and $p^\mu p^\nu p^\alpha p^\beta \to \frac{1}{4!}(\eta^{\mu\nu}\eta^{\alpha\beta}+\eta^{\mu\alpha}\eta^{\nu\beta}+\eta^{\mu\beta}\eta^{\nu\alpha})(p^2)^2$ in the integrals,

we obtain

$$\int \frac{d^4 p}{(2\pi)^4} [\text{tr} \ln M(p)]_4 = -\frac{5}{3} \mathcal{I} \, g^2 f^{abc} f^{abd} \bar{F}^c_{\mu\nu} \bar{F}^{d\mu\nu}$$

$$\int \frac{d^4 p}{(2\pi)^4} [\text{tr} \ln N(p)]_4 = +\frac{1}{12} \mathcal{I} \, g^2 f^{abc} f^{abd} \bar{F}^c_{\mu\nu} \bar{F}^{d\mu\nu} \ .$$

Here we have defined a divergent integral

$$\mathcal{I} \equiv \int \frac{d^4 p}{(2\pi)^4} \frac{1}{(p^2 + i\epsilon)^2} \ .$$

The integral can be performed by computing in $d = 4 - \varepsilon$ dimensions to regulate the UV divergence, and by imposing a lower cutoff m on the magnitude of the (Euclidean) momentum to regulate the IR divergence. After replacing $g \to g\tilde{\mu}^{\varepsilon/2}$ to take care of the mass dimension of the gauge coupling, we have

$$g^2 \mathcal{I} \to i \frac{g^2}{8\pi^2 \varepsilon} \left[1 + \frac{1}{2} \varepsilon \ln\left(\frac{\mu^2}{m^2}\right) + O(\varepsilon^2) \right]$$

where $\mu^2 \equiv 4\pi e^{-(1-\gamma)} \tilde{\mu}^2$ and $\gamma = -\int_0^\infty dt \, e^{-t} \ln t \approx 0.58$.

For the color group $SU(n)$, we have $f^{abc} f^{abd} = n \, \delta^{cd}$ and therefore:

$$\int \frac{d^4 p}{(2\pi)^4} [-\tfrac{1}{2} \text{tr} \ln M(p) + \text{tr} \ln N(p)] = \left(+\frac{5}{6} + \frac{1}{12} \right) \mathcal{I} \, g^2 n \, \bar{F}^a_{\mu\nu} \bar{F}^{a\mu\nu}$$

$$= i \left\{ -\frac{11}{12} n \frac{g^2}{2\pi^2} \left[\frac{1}{\varepsilon} + \frac{1}{2} \ln\left(\frac{\mu^2}{m^2}\right) \right] \right\} \left(-\frac{1}{4} \bar{F}^a_{\mu\nu} \bar{F}^{a\mu\nu} \right)$$

We have therefore deduced the coefficient of $-\frac{1}{4} \bar{F}^a_{\mu\nu} \bar{F}^{a\mu\nu}$ in the 1-loop effective action. Multiplying by $-i$, we find the renormalization factor Z_g to $O(g^2)$:

$$Z_g = 1 - \frac{11}{12} n \frac{g^2}{2\pi^2} \left[\frac{1}{\varepsilon} + \ln\left(\frac{\mu}{m}\right) \right] \ .$$

The bare coupling g_B is related to the running coupling g by $g_B = Z_g^{1/2} g \tilde{\mu}^{\varepsilon/2}$, so that

$$\ln g_B = \mathcal{O}(g, \varepsilon) + \ln g + \tfrac{1}{2} \varepsilon \ln \tilde{\mu} \ .$$

where $\mathcal{O}(g, \varepsilon) \equiv \frac{1}{2} \ln Z_g$. The bare coupling does not know about the parameter μ, so differentiating the above relation gives

$$0 = \frac{\partial \mathcal{O}(g, \varepsilon)}{\partial g} \frac{dg}{d \ln \mu} + \tfrac{1}{2} \varepsilon \ .$$

In general, the function $\mathcal{O}(g, \varepsilon)$ will contain finite terms as well as poles in powers of $1/\varepsilon$:

$$\mathcal{O}(g, \varepsilon) = \mathcal{O}_0(g) + \sum_{n=1}^{\infty} \frac{\mathcal{O}_n(g)}{\varepsilon^n} \ .$$

224

Plugging in this expansion, writing $\lim_{\varepsilon \to 0} \frac{dg}{d\ln\mu} = \beta(g)$ and matching powers of $1/\varepsilon$ we obtain

$$\beta(g) = \tfrac{1}{2} g^2 \mathcal{O}'_1(g) \left[1 - g\mathcal{O}'_0(g) + O\left(g\mathcal{O}'_0(g)\right)^2\right] \; .$$

In our case, we have

$$\mathcal{O}(g,\varepsilon) = \tfrac{1}{2}\ln Z_g = \tfrac{1}{2}\ln\left[1 - \frac{11}{12}n\frac{g^2}{2\pi^2}\left(\frac{1}{\varepsilon} + \ln\frac{\mu}{m}\right)\right] = \frac{11}{12}n\frac{g^2}{4\pi^2}\left(\frac{1}{\varepsilon} + \ln\frac{\mu}{m}\right)$$

where we have dropped terms of $O(g^4)$ and terms of higher powers in $1/\varepsilon$, which are assumed to cancel order by order. Thus matching to the general expansion of $\mathcal{O}(g,\varepsilon)$, we have

$$\mathcal{O}_0(g) = -\frac{11}{12}n\frac{g^2}{4\pi^2}\ln\frac{\mu}{m}$$

$$\mathcal{O}_1(g) = -\frac{11}{12}n\frac{g^2}{4\pi^2} \; .$$

Thus $g\mathcal{O}'_0(g)$ is $O(g^2)$ and can be dropped from the beta function at this order, since it multiplies $g^2\mathcal{O}'_1(g) = O(g^3)$. This is consistent with the general property of gauge theories that 1-loop beta functions are independent of the renormalization scheme.

Since $\mathcal{O}'_1(g) = -\frac{11}{12}n\frac{g}{2\pi^2} = -\frac{11}{3}n\frac{g}{8\pi^2}$, the 1-loop beta function $\beta(g) = \tfrac{1}{2}g^2\mathcal{O}'_1(g)$ is

$$\beta(g) = -\frac{11}{3}n\frac{g^3}{16\pi^2} + O(g^5) \; .$$

This is equation (3) on p. 388. In the non-abelian gauge theory of the strong interaction, we have $n = 3$.

VII.4   Large N

1. Since the number of gluons only differs by one, it is generally argued that it does not make any difference whether we choose to study the $U(N)$ theory or the $SU(N)$ theory. Discuss how the gluon propagator in a $U(N)$ theory differs from the gluon propagator in an $SU(N)$ theory and decide which one is easier.

Solution:

Let $(A_\mu)_i{}^j = \sum_{a=1}^{\dim G} A_\mu^a (T^a)_i{}^j$ be the matrix-valued gluon field for $G = U(N)$ that acts on the defining representation, meaning that the indices i, j run from 1 to N. Then:

$$\langle (A_\mu)_i{}^j(x)(A_\nu)_k{}^\ell(0)\rangle = \langle A_\mu^a(x) A_\nu^b(0)\rangle (T^a)_i{}^j (T^b)_k{}^\ell$$
$$= \Delta_{\mu\nu}(x)\delta^{ab}(T^a)_i{}^j (T^b)_k{}^\ell$$
$$= \Delta_{\mu\nu}(x)(T^a)_i{}^j (T^a)_k{}^\ell$$

where

$$\Delta_{\mu\nu}(x) = \int \frac{d^4k}{(2\pi)^4} \frac{e^{ik\cdot x}}{k^2 + i\varepsilon}\left[-\eta_{\mu\nu} + (1-\xi)\frac{k_\mu k_\nu}{k^2 + i\varepsilon}\right]$$

is the usual abelian massless vector boson propagator in R_ξ gauge. The new feature is just the group theory factor

$$\sum_{a=1}^{N^2}(T^a)_i{}^j (T^a)_k{}^\ell = C\,\delta_i{}^\ell\delta_k{}^j \qquad (1)$$

where C is some constant to be determined as follows. Suppose we normalize the generators as

$$\text{tr}(T^a T^b) = t\,\delta^{ab}$$

where t is a number called the "index" of the defining representation of $U(N)$ to which the generators belong (we have been omitting the subscript N on the generators T^a for typographical convenience.) Setting $b = a$ and summing over a gives $\text{tr}(T^a T^a) = t\delta^{aa} = t\dim U(N) = tN^2$. In (1), setting $k = j$ and summing over j, and setting $\ell = i$ and summing over i gives

$$\text{tr}(T^a T^a) = C\delta_i{}^i\delta_j{}^j = CN^2$$

Therefore $C = t$, which from now on we take equal to 1. The propagator for the $U(N)$ gauge bosons is therefore

$$\langle (A_\mu)_i{}^j(x)(A_\nu)_k{}^\ell(0)\rangle_{U(N)} = \Delta_{\mu\nu}(x)\,\delta_i{}^\ell\delta_k{}^j\,.$$

This provides a simple graphical interpretation, exactly of the form given in figure VII.4.4 on p. 398.

Now suppose we insist on the gauge group $G = SU(N)$ rather than $U(N) = SU(N) \otimes U(1)$. This entails ensuring that the generators are traceless, or in other words insisting on not including the $U(1)$ generator $(T^{a=N^2})_i{}^j = \frac{1}{N^{1/2}}\delta_i{}^j$. All we have to do to modify the propagator

is to take equation (1) (with $C = 1$ as per the previous discussion) and move the $(N^2)^{\text{th}}$ generator from the left-hand side to the right-hand side. In other words,

$$\sum_{a=1}^{N^2-1} (T^a)_i{}^j (T^a)_k{}^\ell = \delta_i{}^\ell \delta_k{}^j - \frac{1}{N} \delta_i{}^j \delta_k{}^\ell \qquad (1')$$

The previous discussion for $U(N)$ carries through exactly, except with the group theory factor from equation $(1')$ instead of (1). Therefore the $SU(N)$ gauge boson propagator is

$$\langle (A_\mu)_i{}^j (x) (A_\nu)_k{}^\ell (0) \rangle_{SU(N)} = \Delta_{\mu\nu}(x) \left(\delta_i{}^\ell \delta_k{}^j - \frac{1}{N} \delta_i{}^j \delta_k{}^\ell \right) .$$

This too has a straightforward graphical interpretation, where we subtract the trace from the $U(N)$ propagator:

2. As a challenge, solve large N QCD in $(1 + 1)$-dimensional spacetime. [Hint: The key is that in $(1 + 1)$-dimensional spacetime with a suitable gauge choice we can integrate out the gauge potential A_μ.] For help, see 't Hooft, *Under the Spell of the Gauge Principle*, p. 443.

Solution:[27]

Define light cone coordinates $p_\pm = \frac{1}{\sqrt{2}}(p_0 \pm p_1)$, so that $2p_+ p_- = p_0^2 - p_1^2 = m^2$, and fix light cone gauge: $A_- = 0$. The Lagrangian for (1+1)-dimensional QCD is then

$$\mathcal{L} = -\tfrac{1}{2} \text{tr}(\partial_- A_+)^2 + \bar{\psi}_i (i \slashed{\partial} - m) \psi^i - g \bar{\psi}_i \gamma_- (A_+)^i{}_j \psi^j$$

where $(A_+)^i{}_j = A_+^a (T^a)^i{}_j$ is the matrix-valued gauge field. Treating x^+ as the temporal direction, we see that A_+ has no dynamics and can be replaced by the 1D Coulomb potential.

The gamma matrices are defined by the Clifford algebra $\{\gamma_\mu, \gamma_\nu\} = 2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, so that $\gamma_+^2 = \gamma_-^2 = 0$ and $\gamma_+ \gamma_- + \gamma_- \gamma_+ = 2$. The only interaction vertex in the theory is $-ig\gamma_-(T^a)^i{}_j$, so the gamma matrices can be removed from the Feynman rules. For example, consider the 1-loop correction to the quark propagator due to single-gluon exchange:

$$-g^2 (T^a T^a)^i{}_j \int \frac{dk_- dk_+}{(2\pi)^2} \frac{\gamma_- \left[\gamma_- (\slashed{k} - \slashed{p})_+ + \gamma_+ (\slashed{k} - \slashed{p})_- + m \right] \gamma_-}{[k_-^2][2(k - p)_+ (k - p)_-]}$$

Since $\gamma_-^2 = 0$ and $\gamma_- \gamma_+ \gamma_- = \gamma_-(-\gamma_- \gamma_+ + 2) = 0 + 2\gamma_-$, only the part of the fermion propagator proportional to γ_+ contributes, and its contribution is simply a factor of 2. Thus the Feynman rules can be taken as:

---

[27]We thank G. 't Hooft for helpful discussion.

$$= \frac{-i}{k_-^2}$$

$$= \frac{ik_-}{2k_+k_- - m^2}$$

$$= -i2g$$

The large-N approximation is to neglect non-planar diagrams and to consider loops due to gluon exchange only. The quark self-energy Σ is related to the exact quark propagator S by

$$iS(k) = \frac{ik_-}{2k_+k_- - m^2 - k_-\Sigma(k)}$$

and satisfies the implicit equation

The above diagram reads

$$i\Sigma(p) = (-2g)^2 \int \frac{dk_- dk_+}{(2\pi)^2} \left(\frac{-i}{k_-^2} \right) \left(\frac{i(k_- + p_-)}{[2(k_+ + p_+) - \Sigma(k+p)](k_- + p_-) - m^2} \right)$$

$$= 4g^2 \int \frac{dk_- dk_+}{(2\pi)^2} \frac{1}{k_-^2} \frac{k_- + p_-}{[2k_+ - \Sigma(k_- + p_-, k_+)](k_- + p_-) - m^2}$$

where in the second line we have shifted $k_+ \to k_+ - p_+$. The integral is independent of p_+, so $\Sigma(p)$ is independent of p_+. We have

$$i\Sigma(p_-) = \frac{4g^2}{(2\pi)^2} \int dk_- \frac{k_- + p_-}{k_-^2} \int dk_+ \frac{1}{2(k_- + p_-)k_+ - (k_- + p_-)\Sigma(k_- + p_-) - m^2} \ .$$

The integral over k_+ is UV divergent. Regularizing via $\int_{-\infty}^{\infty} dk_+ \to \int_{-\Lambda}^{\Lambda} dk_+$, the integral over k_+ evaluates to $\frac{1}{2(k_- + p_-)}\mathrm{sgn}(k_- + p_-)i\pi$. Therefore, we have

$$\Sigma(p_-) = \frac{g^2}{2\pi} \int dk_- \frac{1}{k_-^2} \mathrm{sgn}(k_- + p_-) \ .$$

This integral diverges near $k_- \to 0$. To deal with this, replace $\int_{-\infty}^{\infty} dk_- \to \int_{\mu}^{\infty} dk_- + \int_{-\infty}^{-\mu} dk_-$ with $\mu > 0$ a positive IR regulator. The self-energy is

$$\Sigma(p_-) = \frac{g^2}{2\pi} \left[\int_{\mu}^{\infty} dk_- \frac{1}{k_-^2} \mathrm{sgn}(k_- + p_-) + \int_{-\infty}^{-\mu} dk_- \frac{1}{k_-^2} \mathrm{sgn}(k_- + p_-) \right]$$

$$= \frac{g^2}{2\pi} \int_{\mu}^{\infty} dk_- \frac{1}{k_-^2} [\mathrm{sgn}(p_- + k_-) + \mathrm{sgn}(p_- - k_-)] \ .$$

228

Use $\frac{d}{dx}\text{sgn}(x) = 2\delta(x)$ to evaluate the integral by parts:

$$\Sigma(p_-) = -\frac{g^2}{2\pi} \int_\mu^\infty dk_- \; \frac{d}{dk_-}\left(\frac{1}{k_-}\right) [\text{sgn}(p_- + k_-) + \text{sgn}(p_- - k_-)]$$

$$= -\frac{g^2}{2\pi}\left[\frac{1}{k_-}(\text{sgn}(p_- + k_-) + \text{sgn}(p_- - k_-))\Big|_{k_- = \mu}^{\infty} - \int_\mu^\infty dk_- \frac{2}{k_-}(\delta(p_- + k_-) - \delta(p_- - k_-))\right]$$

$$= -\frac{g^2}{2\pi}\left[-\frac{1}{\mu}(\text{sgn}(p_- + \mu) + \text{sgn}(p_- - \mu)) + 2\text{sgn}(p_-)\frac{1}{|p_-|}\right]$$

$$\to +\frac{g^2}{\pi}\left[\frac{1}{\mu}\text{sgn}(p_-) - \frac{1}{p_-}\right]$$

where in the last line we have taken $\mu \to 0^+$ in the numerator.

Thus we have solved for the quark self-energy. The denominator of the exact propagator is therefore

$$2k_+ k_- - \frac{g^2}{\pi}\left(\frac{|k_-|}{\mu} - 1\right) - m^2 \; .$$

In the limit $\mu \to 0^+$, the pole of the propagator is shifted to $k_+ \to \infty$. This indicates that there is no physical single-quark state.

Next consider the implicit equation given by relating the nth and $(n+1)^{\text{th}}$ ladder diagram:
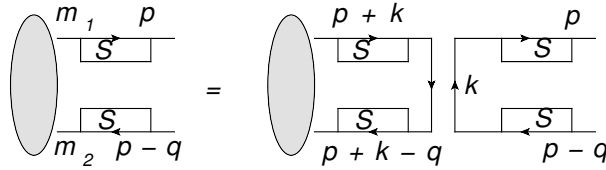
Here the blob, which we denote $\psi(p,q)$, stands for an arbitrary process in which a quark-antiquark pair emerges, the quark with mass m_1 and momentum p, and the antiquark with mass m_2 and momentum $q - p$. The source-free part of this equation is:

$$i\psi(p,q) = (-2g)^2(p_- - q_-)p_i \; \frac{i}{2(p_+ - q_+)(p_- - q_-) - (m_2^2 - \frac{g^2}{\pi}) - \frac{g^2}{\pi\mu}|p_- - q_-|} \times$$

$$\frac{i}{2p_+ p_- - (m_1^2 - \frac{g^2}{\pi}) - \frac{g^2}{\pi\mu}|p_-|} \int \frac{dk_- dk_+}{(2\pi)^2} \frac{-i}{k_-^2} \; i\psi(p+k,q) \; .$$

After a series of manipulations, this can be put into the form of the eigenvalue equation[28]

$$\rho^2 \varphi(x) = \left(\frac{\alpha_1}{x} + \frac{\alpha_2}{1-x}\right)\varphi(x) - \mathcal{P}\int_0^1 dy \; \frac{\varphi(y)}{(y - x)^2}$$

---

[28]The principal value integral is the average value of the integral across the pole:

$$\mathcal{P}\int dx \; \frac{\varphi(x)}{x^2} = \frac{1}{2}\left[\int dx \; \frac{\varphi(x + i\epsilon)}{(x + i\epsilon)^2} + \int dx \; \frac{\varphi(x - i\epsilon)}{(x - i\epsilon)^2}\right] \; .$$

where we have defined $\varphi(p_-, q) \equiv \int dp_+ \psi(p, r)$ and the dimensionless variables $x \equiv p_-/r_-$, $\alpha_i \equiv \frac{\pi}{g^2} m_i^2 - 1$ and $\rho^2 \equiv \frac{\pi}{g^2}(2q_+ q_-)$. Here ρ is the meson mass in units of $g/\pi^{1/2}$.

An approximate solution to the eigenvalue equation may be found by observing that the dominant contribution to the integral on the right-hand side comes from $y \approx x$, at which point the denominator goes to zero. Using the trial function $\varphi(x) = e^{i\omega x}$, we have

$$\mathcal{P} \int_0^1 dy \, \frac{e^{i\omega y}}{(y-x)^2} \approx \mathcal{P} \int_{-\infty}^{\infty} dy \, \frac{e^{i\omega y}}{(y-x)^2} = -\pi|\omega| \, e^{i\omega x}$$

so that $e^{i\omega x}$ is an approximate solution with eigenvalue $\rho^2 \approx \pi|\omega|$, where we have further assumed that the quark and antiquark have equal masses (e.g., a meson made from a $u\bar{u}$ pair rather than a $u\bar{d}$ pair) and satisfy $m_q^2 \approx g^2/\pi$, so that $\alpha_1 = \alpha_2 \approx 0$. The appropriate boundary conditions are $\varphi(0) = \varphi(1) = 0$, so that we are to take the linear combination $\varphi_\omega(x) \propto e^{i\omega x} - e^{-i\omega x} \propto \sin(\omega x)$ as the solution.

Defining $\omega = \pi n$, where n is any positive integer, we thereby arrive at the meson spectrum $m_\pi^2 \equiv \frac{g^2}{\pi} \rho^2$ given by

$$(m_\pi^2)_n \approx g^2 \pi n \, , \quad n = 1, 2, 3, ...$$

The approximation becomes better for larger values of n. In terms of the quark mass m_q, we have

$$\frac{(m_\pi^2)_n}{m_q^2} \approx \pi^2 n \, .$$

For corrections to this lowest order approximation as well as for mesons built of quarks with unequal masses, see the reference [G. 't Hooft, "A two-dimensional model for mesons," Nucl. Phys. B75 (1974) 461-470]. For a path integral treatment as well as for the inclusion of baryons in the spectrum, see E. Witten, "Baryons in the $1/N$ expansion," Nucl. Phys. B160 (1979) 57-115.

3. Show that if we had chosen to calculate $G(z) \equiv \langle (1/N)\text{tr}[1/(z - \varphi)] \rangle$, we would have to connect the two open ends of the quark propagator. We see that figures VII.4.5b and d lead to the same diagram. Complete the calculation of $G(z)$ in this way.

Solution (due to J. Feinberg):

If we want to calculate $G(z)$ directly, then we can't have any free indices dangling on the random matrices φ. The trace takes the upper index on the incoming quark line and contracts it with the lower index on the outgoing quark (which behaves like an incoming antiquark, therefore in an antifundamental rep.) Therefore, we have really closed the loop by connecting the two quark lines in the diagram from the text. We will calculate $G(z)$ by making use of a helpful recursion relation.

As we know from the text, we can write

$$G(z) = \sum_{n=0}^{\infty} \frac{1}{z^{2n+1}} \left\langle \frac{1}{N} \mathrm{tr}\, \varphi^{2n} \right\rangle$$

where the matrix propagator gives a factor of $1/(Nm^2)$. To get the large N limit, the only surviving terms are the planar graphs, which means that the trace, when computed as Wick contractions, breaks into one trace per propagator (plus the original quark line trace). Thus, we can write

$$G(z) = \sum_{n=0}^{\infty} C_n \frac{1}{z^{2n+1}}$$

where C_n is the number of planar terms with n propagators. Let's now derive a recursion relation for the C_n, starting with the zero propagator case, $C_0 = 1$. Let the notation $[\varphi...\varphi]$ denote the canonical Wick contraction, where the two fields at each end (ie, adjacent to each bracket) are contracted. Consider the case with n propagators, or $2n$ matrices. (To get planar graphs, there must be an even number of matrices between each contracted pair.) Thus, the total number of Wick contractions for $2n$ fields is given by

$$(\text{all for } 2n) = [\varphi(\text{all for } 2n-2)\varphi] + [\varphi(\text{all for } 2n-4)\varphi][\varphi\varphi] + ...[\varphi\varphi](\text{all for } 2n-2)$$

$$C_n = C_0 C_{n-1} + C_1 C_{n-2} + ... + C_0 C_{n-1} = \sum_{k=0}^{n-1} C_k C_{n-1-k} \ .$$

Our expression for $G(z)$ becomes

$$G(z) = \frac{1}{z} \sum_{n=0}^{\infty} C_n \left(\frac{1}{z^2 m^2} \right)^n$$

$$= \frac{1}{z} \left[1 + \frac{1}{z^2 m^2} \sum_{n'=0}^{\infty} \sum_{k=0}^{n'} C_k C_{n'-k} \left(\frac{1}{z^2 m^2} \right)^k \left(\frac{1}{z^2 m^2} \right)^{n'-k} \right]$$

$$= \frac{1}{z} \left[1 + \frac{1}{z^2 m^2} \left(\sum_{n=0}^{\infty} C_n \left(\frac{1}{z^2 m^2} \right)^n \right)^2 \right]$$

$$= \frac{1}{z} \left[1 + \frac{1}{m^2} G(z)^2 \right] \ .$$

In the second line, we have used the recursion relation and set $n' = n - 1$. In the third line, we noted that the second line was just the power series expansion of the product of two series. Now we have the same quadratic equation as in the text, leading to

$$G(z) = \frac{m^2}{2} \left(z - \sqrt{z^2 - \frac{4}{m^2}} \right) \ .$$

231

4. Suppose the random matrix φ is real symmetric rather than hermitean. Show that the Feynman rules are more complicated. Calculate the density of eigenvalues. [Hint: The double-line propagator can twist.]

Solution:

First, as clarification, this question is intended to follow the procedure on p. 397 with the quadratic (Gaussian) potential $V(\varphi) = \frac{1}{2}m^2\varphi^2$.

In group theory language, the complication here is that there is no longer any distinction between up and down indices. For φ an $N{\times}N$ hermitian matrix, meaning $(\varphi^\dagger)_i{}^j \equiv (\varphi^j{}_i)^* = \varphi_i{}^j$, we know $\langle\varphi_i{}^j\varphi_k{}^\ell\rangle \propto \delta_i{}^\ell\delta_k{}^j$ by matching upper and lower indices. But if φ is an $N \times N$ real symmetric matrix, meaning $\varphi_{ij} = \varphi_{ji}$ with all entries being real numbers, then matching indices tells us $\langle\varphi_{ij}\varphi_{k\ell}\rangle = C_1\delta_{i\ell}\delta_{jk} + C_2\delta_{ik}\delta_{j\ell}$, with two a priori undetermined constants. The first term corresponds to the hermitian case (just lower the indices ℓ and j from before), while the second term involves a matrix transpose, which diagrammatically appears as a twist in the double-line propagator.

The difference between the complex Hermitian case and the real symmetric case is in calculating the correlation function $\langle(\varphi^2)_{ij}\rangle$. The "propagator" $\langle\varphi_{ik}\varphi_{\ell j}\rangle$ is given by the Gaussian integral

$$\frac{1}{\mathcal{Z}}\int d\varphi\, e^{-N\mathrm{tr}\frac{1}{2}m^2\varphi^2}\varphi_{ik}\varphi_{\ell j} = C_1\delta_{ij}\delta_{k\ell} + C_2\delta_{ik}\delta_{j\ell}$$

where we will now fix the constants C_1 and C_2 by taking special cases of the above expression.

First notice that $\mathrm{tr}\varphi^2 = \varphi_{ij}\varphi_{ji} = \varphi_{11}^2 + 2\varphi_{12}^2 + ...$, meaning that diagonal terms φ_{ii}^2 (no sum) come in with a factor of 1, while off-diagonal terms $\varphi_{i,j\neq i}^2$ come in with a factor of 2.

Let us orient ourselves with a familiar example from ordinary calculus. Define the integral $Z_1 \equiv \int_{-\infty}^{\infty} dx\, e^{-\frac{1}{2}\alpha x^2} = \sqrt{2\pi/\alpha}$. Then for $Z \equiv \prod_{i=1}^{N} Z_1^N$, we have

$$\frac{1}{Z}\int d^N x\, e^{-\frac{1}{2}\alpha\vec{x}^2} x_i x_j = \left[-2\frac{\partial}{\partial\alpha}\ln Z_1\right]\delta_{ij} = +\frac{1}{\alpha}\delta_{ij}\,.$$

Now we will solve for the constants C_1 and C_2 in our matrix theory. First set $i = k = j = \ell = 1$ to get

$$\frac{1}{\mathcal{Z}}\int d\varphi\, e^{-\frac{1}{2}(Nm^2)\varphi_{11}^2+\cdots}\,\varphi_{11}^2 = \frac{1}{\alpha}\bigg|_{\alpha=Nm^2} = \frac{1}{Nm^2} = C_1 + C_2\,.$$

Next set $i = j = 1$, $k = \ell = 2$. Paying special attention to the extra factor of 2 mentioned earlier, we obtain

$$\frac{1}{\mathcal{Z}}\int d\varphi\, e^{-\frac{1}{2}(2Nm^2)\varphi_{12}^2+\cdots}\,\varphi_{12}^2 = \frac{1}{\alpha}\bigg|_{\alpha=2Nm^2} = \frac{1}{2Nm^2} = C_1\,.$$

232

Therefore $C_2 = \frac{1}{Nm^2} - \frac{1}{2Nm^2} = \frac{1}{2Nm^2}$. We arrive at the somewhat intuitively clear result

$$\langle \varphi_{ik}\varphi_{\ell j} \rangle \equiv \frac{1}{\mathcal{Z}} \int d\varphi\, e^{-N\mathrm{tr}\frac{1}{2}m^2\varphi^2} \varphi_{ik}\varphi_{\ell j} = \frac{1}{2Nm^2}\left(\delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell}\right)$$

so that the previous factor $1/(Nm^2)$ is split evenly among the two tensors dictated by group theory.

Setting $k = \ell$ and summing over k gives

$$\langle (\varphi^2)_{ij} \rangle = \frac{1}{2Nm^2}(N\delta_{ij} + \delta_{ik}\delta_{jk}) = \frac{1}{2Nm^2}(N+1)\delta_{ij} \approx \frac{1}{2m^2}\delta_{ij}$$

where we have used the large N limit $N + 1 \approx N$. Thus so far we have computed the $n = 0$ and the $n = 1$ terms in the expansion for $G_{ij}(z)$:

$$G_{ij}(z) = \sum_{n=0}^{\infty} \frac{1}{z^{2n+1}} \langle (\varphi^{2n})_{ij} \rangle = \left(\frac{1}{z} + \frac{1}{z^3}\frac{1}{2m^2} + ... \right) \delta_{ij} \ .$$

As in the text, the next step is to study the planar diagrams contributing to the $n = 2$ term and thereby arrive at a quadratic equation for $G(z)$. The first half of the algorithm ("repeat") translates to the present case exactly as in the text, so that again we arrive at equation (7) on p. 399:

$$G(z) = \frac{1}{z - \Sigma(z)} \ .$$

Next we look at Figure VII.4.6b on p. 400 and observe that in our case we have two terms, one which is depicted in the figure and another one for which the overarching propagator twists. In other words, we have

$$\Sigma(z) = \frac{1}{2m^2}G(z)$$

with the extra factor of $1/2$ arising just as in the above calculation of $\langle (\varphi^2)_{ij} \rangle$. Together these result in the quadratic equation

$$G(z) = \frac{1}{z - G(z)/(2m^2)} \implies G^2 - 2m^2 z\, G + 2m^2 = 0$$

which has the solution

$$G(z) = m^2\left(z - \sqrt{z^2 - \frac{2}{m^2}} \right) \ .$$

This matches equation (9) on p. 400 of the text, rescaled by $m^2 \to 2m^2$. We can understand this result in the following way: Let $\Phi = \frac{1}{\sqrt{2}}(R + iS)$ be a complex matrix field. Then the action $S \equiv m^2 \mathrm{tr}\, \Phi^\dagger\Phi = \frac{1}{2}m^2\mathrm{tr}(R^2 + S^2)$ is the action for two real matrix fields R and S. If Φ is Hermitian, then $\Phi^\dagger = \Phi$ and therefore $S = m^2\mathrm{tr}\, \Phi^\dagger\Phi = m^2\mathrm{tr}\, \Phi^2$. The action in the text is normalized as $\frac{1}{2}m^2\mathrm{tr}\, \Phi^2$, which is brought into the canonical normalization by the rescaling $m^2 \to 2m^2$.

5. For hermitean random matrices φ, calculate

$$G_c(z, w) \equiv \left\langle \frac{1}{N} \text{tr} \frac{1}{z - \varphi} \frac{1}{N} \text{tr} \frac{1}{w - \varphi} \right\rangle - \left\langle \frac{1}{N} \text{tr} \frac{1}{z - \varphi} \right\rangle \left\langle \frac{1}{N} \text{tr} \frac{1}{w - \varphi} \right\rangle$$

for $V(\varphi) = \frac{1}{2} m^2 \varphi^2$ using Feynman diagrams. [Note that this is a much simpler object to study than the object we need to study in order to learn about localization (see exercise VI.6.1).] Show that by taking suitable imaginary parts we can extract the correlation of the density of eigenvalues with itself. For help, see E. Brézin and A. Zee, *Phys. Rev.* E51: p. 5442, 1995.

Solution:

We follow arxiv: cond-mat/9508135v2 9 Oct 1995.

Let $\langle ... \rangle_c$ stand for the connected ensemble average, so that

$$G_c(z, w) = \left\langle \frac{1}{N} \text{tr} \frac{1}{z - \varphi} \frac{1}{N} \text{tr} \frac{1}{w - \varphi} \right\rangle_c$$

$$= \frac{1}{N^2} \frac{1}{zw} \sum_{n,m=1}^{\infty} \left\langle \text{tr} \left(\frac{\varphi}{z} \right)^n \text{tr} \left(\frac{\varphi}{w} \right)^m \right\rangle_c$$

$$= \frac{1}{N^2} \partial_z \partial_w \sum_{n,m=1}^{\infty} \left\langle \text{tr} \left(\frac{\varphi}{z} \right)^n \text{tr} \left(\frac{\varphi}{w} \right)^m \right\rangle_c .$$

In terms of diagrams, each $\text{tr}(\varphi^n)$ is a quark loop, and the average of $\text{tr}(\varphi^n)\text{tr}(\varphi^m)$ tells us that the two quark loops are to be attached to one another through gluon (double-line) propagators.

Since the ensemble is assumed to be gaussian, $V(\varphi) = \frac{1}{2} m^2 \varphi^2$, Wick's theorem tells us that the average $\langle \text{tr}[(\varphi/z)^n] \text{tr}[(\varphi/w)^m] \rangle_c$ is equal to a product of free-field propagators, each consisting of one "z-type" φ and one "w-type" φ, that is one from each trace. Since $\langle \varphi^i{}_j \rangle = 0$, we conclude that only terms for which $n = m$ give a nonzero result:

$$\sum_{n,m=1}^{\infty} \left\langle \text{tr} \left(\frac{\varphi}{z} \right)^n \text{tr} \left(\frac{\varphi}{w} \right)^m \right\rangle_c = \sum_{n=1}^{\infty} \frac{1}{(zw)^n} \langle \text{tr}(\varphi^n)\text{tr}(\varphi^n) \rangle_c .$$

The sum consists of two classes of diagrams, the first in which index contractions are made between the two traces only, and the second in which we allow contractions within the same trace.

For the first class, consider an individual diagram for a fixed value of n, and draw the z-type quark loops inside the w-type quark loops. We have n different ways to attach the z-type loops to the w-type loop (using a double-line propagator connecting from the z-diagram to

one of the loops in the w-diagram). Once this choice is made, there is only one way to attach the rest of the z-loops to the w-loops without having the propagators cross. (Recall we are working in the large-N limit, in which planar diagrams dominate.) Thus each such diagram has a symmetry factor $1/n$.

Given the gluon propagator on p. 398, each attachment of a gluon propagator contributes a factor $1/(Nm^2)$, and each resulting closed loop contributes a factor of N. The factors of N cancel, and we obtain for the sum of these diagrams:

$$
\sum_{n=1}^{\infty} \frac{1}{(zw)^n} \langle \mathrm{tr}(\varphi^n)\mathrm{tr}(\varphi^n) \rangle_c \Big|_{\text{class 1}} = \sum_{n=1}^{\infty} \frac{1}{(zw)^n} \frac{1}{n} \left(\frac{1}{m^2} \right)^n
$$

$$
= -\ln\left(1 - \frac{1}{zwm^2}\right).
$$

Now consider the second class of diagrams, in which we allow contractions within the same trace. In the large-N limit, these diagrams serve only to dress the bare propagator $\frac{1}{z}$ to the full propagator $G(z) = \frac{1}{z-\Sigma(z)}$ (see p. 399). Therefore the full sum of diagrams can be achieved by taking the result from class 1 and simply replacing $1/z$ with $G(z)$ and $1/w$ with $G(w)$. We have

$$
G_c(z,w) = -\frac{1}{N^2} \partial_z \partial_w \ln\left[1 - \frac{1}{m^2} G(z)G(w)\right]
$$

where (as on p. 400)

$$
G(z) = \frac{m^2}{2}\left(z - \sqrt{z^2 - \frac{4}{m^2}}\right).
$$

From this we may obtain the connected correlation between eigenvalues:

$$
\rho_c(E,E') = -\frac{1}{4\pi^2}\left[G_c(+,+) + G_c(-,-) - G_c(+,-) - G_c(-,+)\right]
$$

where $G_c(\pm,\pm) \equiv \lim_{\varepsilon,\varepsilon'\to 0^+} G_c(E \pm i\varepsilon, E' \pm i\varepsilon')$. The result is

$$
\rho_c(E,E') = -\frac{1}{4\pi^2 N^2} \frac{1}{(E-E')^2} \frac{4 - m^2 EE'}{\sqrt{(4-m^2 E^2)(4-m^2 E'^2)}}.
$$

6. Use the Faddeev-Popov method to calculate J in the Dyson gas approach.

Solution:

We will follow footnote 31 on p. 84 of arXiv:hep-th/9304011v1.

The Faddeev-Popov determinant (or Jacobian) J is defined by

$$1 \equiv J(\varphi) \int \mathcal{D}g \, \delta \left(f(\varphi_g) \right) \quad \text{with} \quad \varphi_g = g^\dagger \Lambda g.$$

For the infinitesimal case we have $g = e^{i\theta} = I + i\theta + O(\theta^2)$, so $\varphi_g = \Lambda - i[\theta, \Lambda]$. Moreover, we have for the commutator:

$$[\theta, \Lambda]_{ij} = (\theta\Lambda)_{ij} - (\Lambda\theta)_{ij} = \sum_{k=1}^{N}(\theta_{ik}\lambda_k\delta_{kj} - \lambda_i\delta_{ik}\theta_{kj}) = (\lambda_j - \lambda_i)\theta_{ij}$$

(There are no sums on i and j in the last equality.) So if we choose our gauge transformation $\varphi_g = g^\dagger \Lambda g$ such that $\varphi_g = \Lambda$, or equivalently if we choose the gauge fixing function $f(\varphi) = \varphi - \Lambda$, then the integral over the delta function is:

$$\int \mathcal{D}g \, \delta \left(f(\varphi_g) \right) = \int \mathcal{D}\theta \, \delta(-i[\theta, \Lambda])$$

$$= \prod_{ij} \int d^2\theta_{ij} \, \delta^{(2)} \left((\lambda_j - \lambda_i)\theta_{ij} \right)$$

$$= \prod_{ij} \frac{1}{(\lambda_j - \lambda_i)^2}$$

We wrote $d^2\theta_{ij}$ for each matrix element θ_{ij} because these generator matrices are complex, and we integrate over their real and imaginary parts (or equivalently, over θ_{ij} and θ_{ij}^*). This generates the exponent 2 in the determinant. Therefore, the Faddeev-Popov determinant (Jacobian) is

$$J = \prod_{ij}(\lambda_j - \lambda_i)^2$$

This is the result on p. 401 of the text.

7. For $V(\varphi) = \frac{1}{2}m^2\varphi^2 + g\varphi^4$, determine $\rho(E)$. For m^2 sufficiently negative (the double well potential again) we expect the density of eigenvalues to split into two pieces. This is evident from the Dyson gas picture. Find the critical value m_c^2. For $m^2 < m_c^2$ the assumption of $G(z)$ having only one cut used in the text fails. Show how to calculate $\rho(E)$ in this regime.

Solution:

The function $\sqrt{z^2 - a^2} = \sqrt{(z-a)(z+a)}$ has a single branch cut of finite length along the real axis from $-a$ to $+a$. More generally, the function $\sqrt{(z-a)(z-b)}$ with $b > a$ has a single branch cut of finite length along the real axis from a to b. From this it is clear that the function $\sqrt{(z-a)(z-b)(z+c)(z+d)}$, with a, b, c, d real and positive and $b > a$, $d > c$, has two branch cuts of finite length along the real axis, one from a to b and one from $-d$ to $-c$.

In particular, consider the case $b = ka$, $c = -a$, $d = kc = -ka$ for which the function $f(z) \equiv \sqrt{(z-a)(z-ka)(z+a)(z+ka)} = \sqrt{[z^2 - a^2][z^2 - (ka)^2]}$ contains two disconnected branch cuts of length ka, one from a to ka and the other from $-ka$ to $-a$. We know that the two cuts have the same length by the $\mathbb{Z}_2 : z \to -z$ symmetry of the potential $V(z)$, hence the length ka for both of them. The parameter a fixes how far away from the origin the cuts begin. Together these constitute two unknown parameters.

The function $f(z)$ has the expansion

$$f(z) = z^2 - C - D\frac{1}{z^2} + O\left(\frac{1}{z^4}\right) , \quad C = \frac{1}{2}a^2(k^2 + 1) , \quad D = \frac{1}{8}a^4(k^2 - 1)^2$$

for $z \to \infty$. Following the text, we are motivated to postulate a form

$$G(z) = \frac{1}{2}[V'(z) - P(z)f(z)]$$

where $P(z)$ is a polynomial in z.

Since $V(z) = \frac{1}{2}m^2 z^2 + gz^4 \implies V'(z) = m^2 z + 4gz^3$, we must have $P(z)$ at most linear in z to avoid the introduction of a z^4 term, which cannot be canceled by the cubic polynomial in $V'(z)$. Moreover, $P(z)$ should not have a constant term since that will result in a z^2 term, which also cannot be canceled by the odd polynomial $V'(z)$. Thus we postulate the form $P = Bz$, which introduces a third unknown parameter.

Requiring the cubic and linear terms in $G(z \to \infty)$ to vanish results in two equations. The requirement $G(z \to \infty) \to (+1)\frac{1}{z}$ provides the third equation. The three equations for the three unknowns B, C, D imply

$$B = 4g , \quad C = -4g\,m^2 , \quad D = \frac{1}{2g} .$$

Notice that the definition $C = \frac{1}{2}a^2(k^2+1)$ implies that $C > 0$, so that this solution is only valid for $m^2 < 0$, as expected. Solving for a and k gives two solutions

$$a_\pm = \sqrt{-4g\,m^2 \pm \frac{1}{\sqrt{g}}}\ , \quad k_\pm = \sqrt{\frac{\mp 8g^{3/2}m^2 - (4g^{3/2}m^2)^2 - 1}{1 - (4g^{3/2}m^2)^2}}$$

where the $\pm$ signs are correlated, meaning that (a_+, k_+) is one solution and (a_-, k_-) is the second solution. The solution (a_-, k_-) is only valid when $|m^2| \geq 1/(4g^{3/2})$, which defines the critical value of m^2:

$$-m_c^2 = \frac{1}{4g^{3/2}}$$

For $|m^2| < |m_c^2|$, the two-cut solution ceases to be valid. Let us make sure this is compatible with the solution for k_-. Let $q \equiv 4g^{3/2}m^2 < 0$, so that

$$k_-^2 = \frac{+2q - q^2 - 1}{1 - q^2}\ .$$

The condition $k_-^2 > 0$ is satisfied if $1 - q^2 < 0$, which again implies $|m^2| > 1/(4g^{3/2})$ with $m^2 = -|m^2| < 0$. Thus we have two solutions for the function $G(z)$:

$$G_+(z) = \frac{1}{2}\left[V'(z) - 4gz\sqrt{[z^2 - a_+^2][z^2 - (k_+a_+)^2]}\right]$$

$$G_-(z) = \frac{1}{2}\left[V'(z) - 4gz\sqrt{[z^2 - a_-^2][z^2 - (k_-a_-)^2]}\right]$$

where (a_+, k_+) and (a_-, k_-) are given above as functions of the parameters $m^2 < 0$ and $g > 0$. The density of states $\rho(E) = -\frac{1}{\pi}\text{Im}\,G(z)$ is given by

$$\rho(E) = \begin{cases} 2gE\sqrt{[a_+^2 - E^2][(k_+a_+)^2 - E^2]} & \text{for } a_+ \leq E \leq k_+a_+ \\ 2gE\sqrt{[a_-^2 - E^2][(k_-a_-)^2 - E^2]} & \text{for } a_- \leq E \leq k_-a_- \\ 0 & \text{otherwise} \end{cases}$$

We see that the density of states has split into two disconnected regions.

For a treatment involving orthogonal polynomials, see N. Deo, "Multiple Minima in Glassy Random-Matrix Models," J. Phys.: Condensed Matter, Vol. 12 No. 29, 24 July 2000.

8. Calculate the mass of the soliton (25).

$$m_S = \frac{N}{\pi} m_F \qquad (25)$$

Solution:

After introducing the auxiliary field σ and integrating out the fermions the action is

$$S[\sigma] = -\int d^2x \frac{N}{2g^2}[\sigma(x,t)]^2 - i\tfrac{1}{2}N \operatorname{tr}\ln(-\partial^\mu\partial_\mu - \partial_x\sigma - \sigma)$$

up to an overall additive constant. We know the shape of the time-independent soliton configuration is given by $\sigma(x) = m\tanh(mx)$, where $m = \mu e^{1-\pi/g^2}$, which satisfies the differential equation $\partial_x\sigma + \sigma = m^2$. Thus the entire $\operatorname{tr}\ln(...)$ term is merely an additive constant, and so the energy of the configuration is

$$M = \int_{-\infty}^{\infty} dx \, \frac{N}{2g^2} m^2 \left[C - \tanh^2(mx) \right]$$

where we have written explicitly the additive constant C, which can be thought of as a counterterm to set the vacuum energy to zero. Recall that the coupling constant $g = g(\mu)$ is a function of the renormalization point μ. Choosing $\mu = m$ implies $g^2 = \pi$, so that the energy is

$$M = \frac{Nm}{2\pi} \int_{-\infty}^{\infty} du \, \left[C - \tanh^2 u \right] \, .$$

For a formal derivation of the constant C, see A. Klein, "Bound states and solitons in the Gross-Neveu model," Phys. Rev. D, Vol. 14 No. 2, 15 Jul 1976. Here we will content ourselves with observing that $\lim_{u\to\infty} \tanh u = 1$, so that to get a finite answer we must have $C = 1$. A physical way to see what is going on is to plot the function $1 - \tanh^2 u = \operatorname{sech}^2 u$:
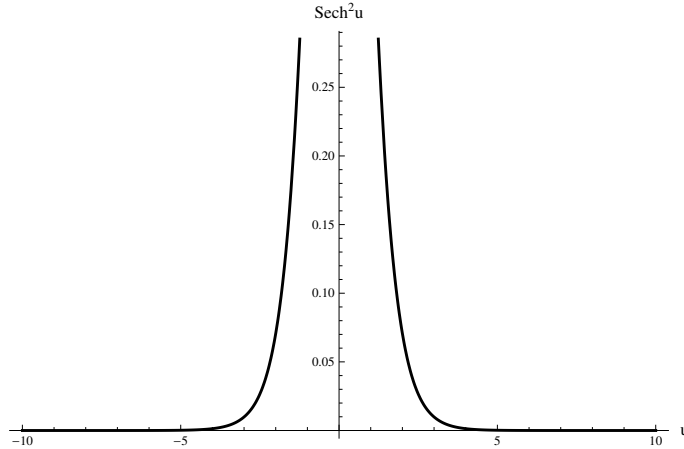
Sech²u

We see that it is a peak localized at the origin, except that it asymptotes to zero only when $C = 1$. The value of the integral is $\int_{-\infty}^{\infty} du \operatorname{sech}^2 u = 2\tanh(\infty) = 2$, so we find

$$M = \frac{Nm}{\pi} \, .$$

239

VII.5   Grand Unification

1. Write down the charge operator Q acting on 5, the defining representation ψ^μ. Work out the charge content of the $10 = \psi^{\mu\nu}$ and identify the various fields contained therein.

Solution:

The generator of electric charge is $Q = T^3 + \frac{1}{2}Y$, where T^3 is the third generator of $SU(2)$. Acting on the $\psi^\mu \sim 5$ of $SU(5)$, we have

$$[Q\psi]^\mu = \left[\begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & \frac{1}{2} & \\ & & & & -\frac{1}{2} \end{pmatrix} + \begin{pmatrix} -\frac{1}{3} & & & & \\ & -\frac{1}{3} & & & \\ & & -\frac{1}{3} & & \\ & & & +\frac{1}{2} & \\ & & & & +\frac{1}{2} \end{pmatrix}\right] \begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \\ \psi^5 \end{pmatrix} = \begin{pmatrix} (-\frac{1}{3})\psi^1 \\ (-\frac{1}{3})\psi^2 \\ (-\frac{1}{3})\psi^3 \\ (+1)\psi^4 \\ (0)\psi^5 \end{pmatrix}$$

Recalling the definition in the text of the $\psi_\mu \sim \bar{5}$ in terms of the familiar fields at low energy

$$\psi_\mu = \begin{pmatrix} \bar{d}_\alpha \\ \nu \\ e \end{pmatrix}$$

you might worry about the charges of the lower two components of the $\psi^\mu \sim 5$ worked out above. When working out the generator of electric charge for the $\bar{5}$, be careful to note that although the $SU(2)$ generator T^3 is the same, the sign of the hypercharge generator flips sign:

$$[Q\psi]_\mu = \left[\begin{pmatrix} 0 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & \frac{1}{2} & \\ & & & & -\frac{1}{2} \end{pmatrix} + \begin{pmatrix} +\frac{1}{3} & & & & \\ & +\frac{1}{3} & & & \\ & & +\frac{1}{3} & & \\ & & & -\frac{1}{2} & \\ & & & & -\frac{1}{2} \end{pmatrix}\right] \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \\ \psi_5 \end{pmatrix} = \begin{pmatrix} (+\frac{1}{3})\psi_1 \\ (+\frac{1}{3})\psi_2 \\ (+\frac{1}{3})\psi_3 \\ (0)\psi_4 \\ (-1)\psi_5 \end{pmatrix}$$

So the electron field $e = \psi_5$ indeed has charge -1, and the neutrino field $\nu = \psi_4$ is neutral.

Now for the 10. Given a representation R and its generators T_R, and given a second representation R' and its generators T'_R, the generators of the product representation $R \otimes R'$ are given by $T_{R \otimes R'} = T_R \otimes I_{R'} + I_R \otimes T_{R'}$. Explicitly in terms of components, if $\mu, \nu, ...$ denote the matrix indices of representation R and if $a, b, ...$ denote the matrix indices of representation R', then

$$(T_{R \otimes R'})^{\mu a}{}_{\nu b} = (T_R)^\mu{}_\nu \delta^a{}_b + \delta^\mu{}_\nu (T_{R'})^a{}_b \, .$$

If the representations R' and R are the same, then we can symmetrize and antisymmetrize the two upper indices (or the two lower indices). The generator of the symmetric product representation $R \otimes_S R$ is

$$(T_{R \otimes_S R})^{\mu\nu}{}_{\rho\sigma} = (T_R)^{(\mu}{}_\rho \delta^{\nu)}{}_\sigma + \delta^{(\mu}{}_\rho (T_R)^{\nu)}{}_\sigma$$

240

where the parentheses mean symmetrization in those indices:

$$M^{(\mu\nu)} \equiv \frac{1}{2}(M^{\mu\nu} + M^{\nu\mu}).$$

Similarly, the generator of the antisymmetric product representation $R \otimes_A R$ is

$$(T_{R \otimes_A R})^{\mu\nu}{}_{\rho\sigma} = (T_R)^{[\mu}{}_{\rho}\delta^{\nu]}{}_{\sigma} + \delta^{[\mu}{}_{\rho}(T_R)^{\nu]}{}_{\sigma}$$

where the brackets mean antisymmetrization in those indices:

$$M^{[\mu\nu]} \equiv \frac{1}{2}(M^{\mu\nu} - M^{\nu\mu}).$$

The 10 of $SU(5)$ is the antisymmetric product representation $5 \otimes_A 5$, so the generator of electric charge acting on the 10 representation is

$$(Q_{10})^{\mu\nu}{}_{\rho\sigma} = (Q_5)^{[\mu}{}_{\rho}\delta^{\nu]}{}_{\sigma} + \delta^{[\mu}{}_{\rho}(Q_5)^{\nu]}{}_{\sigma}$$

where $Q_5 = \text{diag}(-\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}, +1, 0)$ is the generator of electric charge acting on the 5 of $SU(5)$. Defining $\psi'^{\mu\nu} = (Q_{10})^{\mu\nu}{}_{\rho\sigma}\psi^{\rho\sigma}$, we find

$$\psi'^{\alpha\beta} = (-\frac{2}{3})\psi^{\alpha\beta} \implies \psi^{\alpha\beta} = \varepsilon^{\alpha\beta\gamma}\bar{u}_{\gamma}$$
$$\psi'^{45} = (+1)\psi^{45} \implies \psi^{45} = \bar{e} \ .$$

We also have $\psi'^{\alpha i} = (Q_{10})^{\alpha i}{}_{\beta j}\psi^{\beta j} = (-\frac{1}{3}\delta^i{}_j + \delta^i{}_4\delta^4{}_j)\psi^{\alpha j}$, so:

$$\psi'^{\alpha 4} = (-\frac{1}{3} + 1)\psi^{\alpha 4} = +\frac{2}{3}\psi^{\alpha 4} \implies \psi^{\alpha 4} = u^{\alpha}$$
$$\psi'^{\alpha 5} = (-\frac{1}{3} + 0)\psi^{\alpha 5} = -\frac{1}{3}\psi^{\alpha 5} \implies \psi^{\alpha 5} = d^{\alpha} \ .$$

Actually there is a slight mistake here. The lepton doublet $\ell_i \equiv \begin{pmatrix} \nu \\ e \end{pmatrix}$ is defined correctly with a lower $SU(2)$ index, since it is part of the $\bar{5}$ $\psi_{\mu} = \begin{pmatrix} d^c_{\alpha} \\ \ell_i \end{pmatrix}$. The usual Standard-Model convention for the quark doublet is $q_i \equiv \begin{pmatrix} u \\ d \end{pmatrix}$ with a lower $SU(2)$ index just like for the leptons. This implies $q^i = \varepsilon^{ij}q_j = \begin{pmatrix} d \\ -u \end{pmatrix}$, so that $\psi^{\alpha 5} = d^{\alpha}$ as written, but $\psi^{\alpha 4} = -u^{\alpha}$ with an extra minus sign arising from $\varepsilon^{54} = -\varepsilon^{45}$.

2. Show that for any grand unified theory, as long as it is based on a simple group, we have at the unification scale

$$\sin^2 \theta = \frac{\sum T_3^2}{\sum Q^2}$$

where the sum is taken over all fermions.

Solution:

Consider the case for one fermion ψ that transforms under some representation of the simple group G. Its gauge-covariant derivative in the electroweak subgroup $SU(2) \otimes U(1)$ is

$$D_\mu \psi = \left[\partial_\mu - ig \left(A_\mu^a T_a + B_\mu \sqrt{\alpha} \frac{1}{2} Y \right) \right] \psi$$

As discussed on p. 410 of the text for the particular case $\alpha = 3/5$, we fix α by demanding that the $SU(2)$ and $U(1)$ generators are normalized equally, or in other words $\text{tr}[(T_3)^2] = \text{tr}[(\sqrt{\alpha} \frac{1}{2} Y)^2]$. This implies

$$\alpha = \frac{\text{tr}[(T_3)^2]}{\text{tr}[(\frac{1}{2} Y)^2]}$$

We define the weak mixing angle by $\tan \theta \equiv g_1/g_2 = \sqrt{\alpha}$, so since $\tan \theta = s/c = s/\sqrt{1 - s^2}$, where $c \equiv \cos \theta$ and $s \equiv \sin \theta$, we have

$$s^2 = \frac{\alpha}{1 + \alpha} \ .$$

The generator of electric charge is $Q = T_3 + \frac{1}{2} Y$, so $\text{tr}[(\frac{1}{2} Y)^2] = \text{tr}(Q^2 - 2QT_3 + T_3^2)$ and therefore

$$s^2 = \frac{\text{tr}(T_3^2)}{2 \, \text{tr}[T_3(T_3 - Q)] + \text{tr}(Q^2)} \ .$$

But $T_3 - Q = -\frac{1}{2} Y$, and $\text{tr}(T_3 \frac{1}{2} Y) = 0$. Therefore, for a single fermion, we have

$$s^2 = \frac{\text{tr}(T_3^2)}{\text{tr}(Q^2)}$$

at the unification scale.

In the case of multiple fermions, the currents in each direction of the Lie algebra get contributions from each fermion f that transforms under the group. In other words, we fix α by $\sum_f \text{tr}[(T_3^{(f)})^2] = \sum_f \text{tr}[(\sqrt{\alpha} \frac{1}{2} Y^{(f)})^2]$, where $T_3^{(f)}$ and $\frac{1}{2} Y^{(f)}$ are the generators appropriate for the representation to which fermion f belongs. Everything carries through as before, and we arrive at the result

$$s^2 = \frac{\sum_f \text{tr}[(T_3^{(f)})^2]}{\sum_f \text{tr}[(Q^{(f)})^2]} \ .$$

242

3. Check that the $SU(3) \otimes SU(2) \otimes U(1)$ theory is anomaly-free. [Hint: The calculation is more involved than in $SU(5)$ since there are more independent generators. First show that you only have to evaluate tr $Y[T_a, T_b]$ and tr Y^3, with T_a and Y the generators of $SU(2)$ and $U(1)$, respectively.]

Solution (due to J. Feinberg):

Actually, we have to calculate slightly more than what is stated in the original problem. In addition to denoting $SU(2)$ generators by T, denote normalized $SU(3)$ generators as t. One thing we can note: since the charges are repeated in each generation of particles, we only need to calculate the anomaly cancellation for one of them. Note that when we say "tr" here, we mean to treat the gauge groups as tensor products (so the traces factorize) and to put an additional minus sign on the right-handed field traces.

We start with anomalies involving $SU(3)$. Since $SU(3)$ couples equally to left and right handed particles, the one with 3 ts vanishes automatically. With only one t, one of the factors in the trace of generators is tr $t = 0$, so that always vanishes also. Now consider the traces with 2 ts. We have $\mathrm{tr}(t^a t^b)\mathrm{tr}T^c = 0$ and $\mathrm{tr}(t^a t^b)\mathrm{tr}Y = \frac{1}{2}\delta^{ab}\mathrm{tr}_q Y$. The trace now runs only over the quarks because the leptons have $t = 0$. We have $\mathrm{tr}_q Y = 2 \times (\frac{1}{6}) - \frac{2}{3} - (-\frac{1}{3}) = 0$. Note that the factor of 2 in the first term is because the left-handed quarks are in an $SU(2)$ doublet, and that we have subtracted the right-handed ones.

Now take anomalies involving $SU(2)$ generators. Since the fermions are all in doublets or singlets, we can use the Pauli matrix anticommutator, which gives $\{T^a, T^b\} = \frac{1}{2}\delta^{ab}$. Thus, the $3T$ anomaly is $\mathrm{tr}(\{T^a, T^b\}T^c) = \frac{1}{2}\delta^{ab}\mathrm{tr}(T^c) = 0$. Similarly, $\mathrm{tr}(Y^2 T^a) = \mathrm{tr}Y^2\mathrm{tr}T^c = 0$. The last case is with 2 Ts, $\mathrm{tr}(YT^aT^b) = \frac{1}{2}\delta^{ab}\mathrm{tr}_L Y$, where the trace is now over left-handed fields. This gives $\mathrm{tr}_L Y = 3(\frac{1}{6}) - \frac{1}{2} = 0$. Note the factor of 3 because the quarks are in an $SU(3)$ triplet.

Finally, we have traces involving only hypercharge. This is $\mathrm{tr}Y^3 = 3 \times 2 \times (\frac{1}{6}) + 2(-\frac{1}{2}) - 3(\frac{2}{3}) - 3(-\frac{1}{3}) - (-1) = 0$. Technically, one should also calculate gravitational couplings, but we are not considering that in this model.

4. Construct grand unified theories based on $SU(6)$, $SU(7)$, $SU(8)$, ... , until you get tired of the game. People used to get tenure doing this. [Hint: You would have to invent fermions yet to be experimentally discovered.]

Solution:

This is of course an open-ended problem with many solutions. We will simply point towards some of the literature on unified models based on $SU(n)$ for $n > 5$:

- $SU(6)$: M. Singer and K. S. Viswanathan, "SU(6) Grand Unified Theory without Fundamental Scalars," Phys. Rev. D Vol. 24 No. 11, 1 Dec 1981

- $SU(7)$: H. Goldberg, T. W. Kephart and M. T. Vaughn, "Fractionally Charged Color-Singlet Fermions in a Grand Unified Theory," Phys. Rev. Lett. Vol. 47, No. 20, 16 Nov. 1981

- $SU(9)$: Y. Fujimoto and P. Sodano, "SU(9) Grand Unified Theory," Phys. Rev. D Vol. 23 No. 7, 1 Apr. 1981

- $SU(11)$: H. Georgi, "Towards a Grand Unified Theory of Flavor," Nucl. Phys. B, 1979

- $SU(15)$: P. Frampton and B. H. Lee, "SU(15) Grand Unification," Phys. Rev. Lett. Vol. 64 No. 6, 5 Feb 1990

VII.6   Protons Are Not Forever

1. Suppose there are F' new families of quarks and leptons with masses of order M'. Adopting the crude approximation described in exercise IV.8.2 of ignoring these families for μ below M' and of treating M' as negligible for μ above M', run the renormalization group flow and discuss how various predictions, such as proton lifetime, are changed.

Solution:

For energies below the scale M' we still have only F "massless" fermion families, so equations (1)-(3) on p. 414 become

$$\frac{1}{\alpha_S(\mu)} = \frac{1}{\alpha_S(M')} + \frac{1}{6\pi}(4F - 33)\ln\left(\frac{M'}{\mu}\right)$$

$$\frac{\sin^2\theta(\mu)}{\alpha(\mu)} = \frac{\sin^2\theta(M')}{\alpha(M')} + \frac{1}{6\pi}(4F - 22)\ln\left(\frac{M'}{\mu}\right)$$

$$\frac{\cos^2\theta(\mu)}{\alpha(\mu)} = \frac{\cos^2\theta(M')}{\alpha(M')} + \frac{1}{6\pi}\frac{20}{3}F\ln\left(\frac{M'}{\mu}\right) .$$

Above the scale M' we excite the new degrees of freedom, so for $M' < \mu < M_{\text{GUT}}$, these equations become exactly those on p. 414 except with F replaced by $F + F'$:

$$\frac{1}{\alpha_S(\mu)} = \frac{1}{\alpha_S(M_{\text{GUT}})} + \frac{1}{6\pi}[4(F + F') - 33]\ln\left(\frac{M_{\text{GUT}}}{\mu}\right)$$

$$\frac{\sin^2\theta(\mu)}{\alpha(\mu)} = \frac{\sin^2\theta(M_{\text{GUT}})}{\alpha(M_{\text{GUT}})} + \frac{1}{6\pi}[4(F + F') - 22]\ln\left(\frac{M_{\text{GUT}}}{\mu}\right)$$

$$\frac{\cos^2\theta(\mu)}{\alpha(\mu)} = \frac{\cos^2\theta(M_{\text{GUT}})}{\alpha(M_{\text{GUT}})} + \frac{1}{6\pi}\frac{20}{3}(F + F')\ln\left(\frac{M_{\text{GUT}}}{\mu}\right) .$$

As discussed in the text, the couplings are assumed to unify at the scale of grand unification $\alpha(M_{\text{GUT}}) = \alpha_S(M_{\text{GUT}}) \equiv \alpha_{\text{GUT}}$, and the angles are normalized by $\tan^2\theta(M_{\text{GUT}}) = 3/5$.

The best data we have is provided at the mass of the Z boson, $M_Z = 91.188$ GeV. We use the first set of equations to specify $\alpha_S(M')$, $\alpha(M')$ and $\sin^2\theta(M')$ in terms of the parameters evaluated at $\mu = M_Z$. We then use these values of the parameters at M' as inputs into the second set of equations, which run up to M_{GUT}. This results in

$$\frac{1}{\alpha_S(M_{\text{GUT}})} = \frac{1}{\alpha_S(M_Z)} - \frac{1}{6\pi}(4F - 33)\ln\left(\frac{M_{\text{GUT}}}{M_Z}\right) - \frac{1}{6\pi}4F'\ln\left(\frac{M_{\text{GUT}}}{M'}\right)$$

$$\frac{\sin^2\theta(M_{\text{GUT}})}{\alpha(M_{\text{GUT}})} = \frac{\sin^2\theta(M_Z)}{\alpha(M_Z)} - \frac{1}{6\pi}(4F - 22)\ln\left(\frac{M_{\text{GUT}}}{M_Z}\right) - \frac{1}{6\pi}4F'\ln\left(\frac{M_{\text{GUT}}}{M'}\right)$$

$$\frac{\cos^2\theta(M_{\text{GUT}})}{\alpha(M_{\text{GUT}})} = \frac{\cos^2\theta(M_Z)}{\alpha(M_Z)} - \frac{1}{6\pi}\frac{20}{3}F\ln\left(\frac{M_{\text{GUT}}}{M_Z}\right) - \frac{1}{6\pi}\frac{20}{3}F'\ln\left(\frac{M_{\text{GUT}}}{M'}\right) .$$

You can see what is happening: if $F' = 0$, then we recover the equations on p. 414 with $\mu = M_Z$. The presence of the new families when $F' \neq 0$ affects the running between the

245

scales M' and M_{GUT}.

See W. J. Marciano, "Weak mixing angle and grand unified gauge theories," Phys. Rev. D, Vol. 20 No. 1, 1 July 1979 for running the renormalization group flow to compare $SU(5)$ GUT predictions with experiment.

2. Work out proton decay in detail. Derive relations between the following decay rates: $\Gamma(p \to \pi^0 e^+)$, $\Gamma(p \to \pi^+ \bar\nu)$, $\Gamma(n \to \pi^- e^+)$, and $\Gamma(n \to \pi^0 \bar\nu)$.

Solution:

We will use two-component spinor notation (see Appendix E). As explained on p. 407 in the text, we write all fields as left-handed with the following transformation properties under $G_{\text{SM}} \equiv SU(3)_{\text{c}} \otimes SU(2)_{\text{W}} \otimes U(1)_{\text{Y}}$:

$$q \equiv \begin{pmatrix} u \\ d \end{pmatrix} \sim (3, 2, \tfrac{1}{6}),\ \bar{u} \sim (\bar{3}, 1, -\tfrac{2}{3}),\ \bar{d} \sim (\bar{3}, 1, \tfrac{1}{3}),\ \ell \equiv \begin{pmatrix} \nu \\ e \end{pmatrix} \sim (1, 2, -\tfrac{1}{2}),\ \bar{e} \sim (1, 1, 1)$$

The indices α, β, γ run from 1 to 3 and denote the 3-representation of $SU(3)_{\text{c}}$, and the indices i, j run from 1 to 2 and denote the 2-representation of $SU(2)_{\text{W}}$. The antisymmetric tensor $\varepsilon^{\alpha\beta\gamma}$ is invariant under $SU(3)_{\text{c}}$ and the antisymmetric tensor ε^{ij} is invariant under $SU(2)_{\text{W}}$.

The $SU(5)$ fermion fields are $(\psi_{\bar5})_\mu \equiv \begin{pmatrix} \bar{d}_\alpha \\ \ell_i \end{pmatrix}$ and $(\psi_{10})^{\mu\nu} = -(\psi_{10})^{\nu\mu}$, whose components are

$(\psi_{10})^{\alpha\beta} = \varepsilon^{\alpha\beta\gamma}\bar{u}_\gamma$, $(\psi_{10})^\alpha_i = q^\alpha_i = \begin{pmatrix} u^\alpha \\ d^\alpha \end{pmatrix}$ and $(\psi_{10})^{ij} = \varepsilon^{ij}\bar{e}$. As pointed out in problem VII.5.1,

we have $(\psi_{10})^{\alpha i} = \varepsilon^{ij}(\psi_{10})^\alpha_j$ so that $(\psi_{10})^{\alpha 4} = \varepsilon^{45}d^\alpha = +d^\alpha$ and $(\psi_{10})^{\alpha 5} = \varepsilon^{54}u^\alpha = -u^\alpha$.

Let us now work out the new currents that arise from the $SU(5)$ unified theory. The covariant derivative acting on a 5 of $SU(5)$ is

$$(D\psi_5)^\mu = \partial\psi_5^\mu - ig\, X^\mu{}_\nu \psi_5^\nu\,, \quad X^\mu{}_\nu = \sum_{a=1}^{24} X^a (T_5^a)^\mu{}_\nu\,.$$

At this point we should note an annoying circumstance: Since we have already used $\mu = 1, ..., 5$ as an index for $SU(5)$, $\alpha = 1, 2, 3$ as an index for $SU(3)$, $i = 1, 2$ as an index for $SU(2)$ and $a = 1, ..., 24$ as an index for the adjoint representation of $SU(5)$, we will denote Lorentz vector indices by $M = 0, 1, 2, 3$, and Lorentz spinor indices by $m = 1, 2$ and $\dot{m} = 1, 2$ when we choose to display them explicitly.

The covariant derivative acting on a $\bar{5}$ of $SU(5)$ is

$$(D\psi_{\bar5})_\mu = \partial\psi_{\bar5\,\mu} - ig\, \bar{X}_\mu{}^\nu \psi_{\bar5\,\nu}\,, \quad \bar{X}_\mu{}^\nu = \sum_{a=1}^{24} X^a (T_{\bar5}^a)_\mu{}^\nu\,.$$

246

The adjoint components X^a are the same for the 5 and for the $\bar{5}$. The generators of the $\bar{5}$ are $T_{\bar{5}}^a = -(T_5^a)^* = -(T_5^a)^T$ (since T_5^a are hermitian, we have $(T_5^a)^* = (T_5^a)^T$). In components, this means $(T_{\bar{5}}^a)_\mu{}^\nu = -[(T_5^a)^T]_\mu{}^\nu = -(T_5^a)^\nu{}_\mu$. Therefore the covariant derivative of a $\bar{5}$ is

$$(D\psi_{\bar{5}})_\mu = \partial\psi_{\bar{5}\mu} + ig \sum_{a=1}^{24} X^a (T_5^a)^\nu{}_\mu \psi_{\bar{5}\nu}$$

so the kinetic term in the Lagrangian for $\psi_{\bar{5}}$ is

$$\mathcal{L} = i\psi_{\bar{5}}^\dagger D\psi_{\bar{5}}$$
$$= i\psi_{\bar{5}}^\dagger \partial\psi_{\bar{5}} - g\sum_{a=1}^{24} \psi_{\bar{5}}^{\dagger\,\mu} X^a (T_5^a)^\nu{}_\mu \psi_{\bar{5}\nu} \ .$$

This form for the current term is awkward because the $SU(5)$ indices are not in matrix multiplication order. To take care of this properly, we should display the Lorentz indices:

$$\psi_{\bar{5}}^{\dagger\,\mu} X^a (T_5^a)^\nu{}_\mu \psi_{\bar{5}\nu} = X_M^a (\psi_{\bar{5}}^{\dagger\,\mu})_{\dot{m}} \bar{\sigma}^{M\dot{m}m} (T_5^a)^\nu{}_\mu (\psi_{\bar{5}\nu})_m$$
$$= -X_M^a (\psi_{\bar{5}\nu})^m \sigma_{m\dot{m}}^M (T_5^a)^\nu{}_\mu (\psi_{\bar{5}}^{\dagger\,\mu})^m$$
$$= -\psi_{\bar{5}\nu} X^a (T_5^a)^\nu{}_\mu \psi_{\bar{5}}^{\dagger\,\mu}$$

Now the $SU(5)$ indices are in matrix multiplication order. The kinetic term in the Lagrangian for $\psi_{\bar{5}}$ is

$$\mathcal{L} = i\psi_{\bar{5}}^\dagger \partial\psi_{\bar{5}} + g\psi_{\bar{5}} X \psi_{\bar{5}}^\dagger$$

where $X = \sum_{a=1}^{24} X^a$, and now all indices can be suppressed without any confusion. The idea now is to see what this current implies for the low-energy fields $\bar{d}$ and ℓ. Since

$$\psi_{\bar{5}\mu} X^\mu{}_\nu \psi_{\bar{5}}^{\dagger\,\nu} = \psi_{\bar{5}\alpha} X^\alpha{}_\beta \psi_{\bar{5}}^{\dagger\,\beta} + \psi_{\bar{5}\alpha} X^\alpha{}_i \psi_{\bar{5}}^{\dagger\,i} + \psi_{\bar{5}i} X^i{}_\alpha \psi_{\bar{5}}^{\dagger\,\alpha} + \psi_{\bar{5}i} X^i{}_j \psi_{\bar{5}}^{\dagger\,j}$$

and $\psi_{\bar{5}\alpha} = \bar{d}_\alpha$, $\psi_{\bar{5}i} = \ell_i$, we can identify the new baryon-number-violating interactions from the $\bar{5}$ as

$$\mathcal{L}_{\bar{5}}^{\Delta B \neq 0} = g\left(\bar{d}_\alpha X^\alpha{}_i \ell^{\dagger i} + h.c.\right)$$

Now for the 10 of $SU(5)$. The covariant derivative of ψ_{10} is

$$(D\psi_{10})^{\mu\nu} = \partial\psi_{10}^{\mu\nu} - ig\sum_{a=1}^{24} X^a (T_{10}^a)^{\mu\nu}{}_{\rho\sigma} \psi_{10}^{\rho\sigma}$$

where as discussed in problem VII.5.1 the generators of the $10 = 5 \otimes_A 5$ are

$$(T_{10}^a)^{\mu\nu}{}_{\rho\sigma} = (T_5^a)^{[\mu}{}_\rho \delta^{\nu]}{}_\sigma + \delta^{[\mu}{}_\rho (T_5^a)^{\nu]}{}_\sigma$$

where we remind you of the notation $f^{[\mu\nu]} \equiv \frac{1}{2}(f^{\mu\nu} - f^{\nu\mu})$. From the kinetic term $\mathcal{L} = \frac{1}{2}i\psi_{10}^\dagger D\psi_{10}$ we therefore get the baryon-number-violating interactions

$$\mathcal{L}_{10}^{\Delta B \neq 0} = g[(\psi_{10}^\dagger)_{\alpha\beta} X^\alpha{}_i \psi^{i\beta} + (\psi_{10}^\dagger)_{\alpha j} X^\alpha{}_i \psi^{ij} + h.c.]$$
$$= g[\varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} X^\alpha{}_i (-\varepsilon^{ij} q_j^\beta) + (\varepsilon_{jk} q_\alpha^{\dagger k}) X^\alpha{}_i \varepsilon^{ij} \bar{e} + h.c.]$$
$$= g[-\varepsilon^{ij} \varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} X^\alpha{}_i q_j^\beta + q_\alpha^{\dagger i} X^\alpha{}_i \bar{e} + h.c.]$$

247

Putting together the contributions from the $\bar{5}$ and the 10, we have

$$\mathcal{L}^{\Delta B \neq 0} = g(X_M)_i^\alpha \left(\bar{d}_\alpha \sigma^M \ell^{\dagger i} - \varepsilon^{ij} \varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} \bar{\sigma}^M q_j^\beta + q_\alpha^{\dagger i} \bar{\sigma}^M \bar{e} + h.c. \right)$$

Expanding out the $SU(2)$ index with $\ell_i = \begin{pmatrix} \nu \\ e \end{pmatrix}$, we arrive at the interactions

$$\mathcal{L}^{\Delta B \neq 0} = g(X_M)_4^\alpha \left(\bar{d}_\alpha \sigma^M \nu^\dagger - \varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} \bar{\sigma}^M d^\beta + u_\alpha^\dagger \bar{\sigma}^M \bar{e} + h.c. \right)$$
$$+ g(X_M)_5^\alpha \left(\bar{d}_\alpha \sigma^M e^\dagger + \varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} \bar{\sigma}^M u^\beta + d_\alpha^\dagger \bar{\sigma}^M \bar{e} + h.c. \right)$$

Recall the electric charge of the down quark $Q(d) = -\frac{1}{3}$. Since $Q(\bar{d}\nu^\dagger) = -Q(d) + Q(\nu) = +\frac{1}{3} + 0$, invariance under electromagnetism tells us that the gauge boson $(X_M)_4^\alpha$ has electric charge $-\frac{1}{3}$. Since $Q(\bar{d}e^\dagger) = -Q(d) - Q(e) = +\frac{1}{3} + 1 = +\frac{4}{3}$, the gauge boson $(X_M)_5^\alpha$ has electric charge $-\frac{4}{3}$. Let us therefore write $(X^{-1/3})_M^\alpha \equiv (X_M)_4^\alpha$ and $(X^{-4/3})_M^\alpha \equiv (X_M)_5^\alpha$ and rewrite the above interactions as

$$\mathcal{L}^{\Delta B \neq 0} = g\,(X^{-1/3})_M^\alpha\,(J^{+1/3})_\alpha^M + g\,(X^{-4/3})_M^\alpha\,(J^{+4/3})_\alpha^M + h.c.$$

where we have defined the electrically charged currents

$$(J^{+1/3})_\alpha^M \equiv \bar{d}_\alpha \sigma^M \nu^\dagger - \varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} \bar{\sigma}^M d^\beta + u_\alpha^\dagger \bar{\sigma}^M \bar{e}$$
$$(J^{+4/3})_\alpha^M \equiv \bar{d}_\alpha \sigma^M e^\dagger + \varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} \bar{\sigma}^M u^\beta + d_\alpha^\dagger \bar{\sigma}^M \bar{e} .$$

As a consistency check, we should verify the electric charges of all of the terms in the currents. We have $Q(\bar{u}^\dagger d) = +Q(u) + Q(d) = +\frac{2}{3} - \frac{1}{3} = +\frac{1}{3}$ and $Q(u^\dagger \bar{e}) = -Q(u) - Q(e) = -\frac{2}{3} + 1 = +\frac{1}{3}$, which both match $Q(\bar{d}\nu^\dagger) = +\frac{1}{3}$. Also $Q(\bar{u}^\dagger u) = 2Q(u) = +\frac{4}{3}$ and $Q(d^\dagger \bar{e}) = -Q(d) - Q(e) = +\frac{1}{3} + 1 = +\frac{4}{3}$, which match $Q(\bar{d}e^\dagger) = +\frac{4}{3}$ as they must.

Suppose the gauge bosons $X^{-1/3}$ and $X^{-4/3}$ have the same GUT-scale mass $M_X \sim 10^{16}$ GeV so that they can be integrated out at low energies. Performing this integration gives the low-energy effective Lagrangian for baryon-number-violating processes such as proton decay:

$$\mathcal{L}_{\text{eff}}^{\Delta B \neq 0} = \frac{g^2}{M_X^2} \left[(J^{+1/3})_\alpha^M (J^{-1/3})_M^\alpha + (J^{+4/3})_\alpha^M (J^{-4/3})_M^\alpha \right] .$$

The products of currents in this Lagrangian are

$$J^{+1/3} J^{-1/3} = (\bar{d}_\alpha \sigma^M \nu^\dagger - \varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} \bar{\sigma}^M d^\beta + u_\alpha^\dagger \bar{\sigma}^M \bar{e})(\nu \sigma_M \bar{d}^{\dagger\alpha} - \varepsilon^{\alpha\delta\epsilon} d_\delta^\dagger \bar{\sigma}_M \bar{u}_\epsilon + \bar{e}^\dagger \bar{\sigma}_M u^\alpha)$$
$$J^{+4/3} J^{-4/3} = (\bar{d}_\alpha \sigma^M e^\dagger + \varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} \bar{\sigma}^M u^\beta + d_\alpha^\dagger \bar{\sigma}^M \bar{e})(e \sigma_M \bar{d}^{\dagger\alpha} + \varepsilon^{\alpha\delta\epsilon} u_\delta^\dagger \bar{\sigma}_M \bar{u}_\epsilon + \bar{e}^\dagger \bar{\sigma}_M d^\alpha)$$

For this problem, we are interested in only those interactions that contribute to proton decay and neutron decay, so we are interested in operators of the schematic form $\sim qqq\ell$. These are:

$$J^{+1/3} J^{-1/3} = (\bar{d}_\alpha \sigma^M \nu^\dagger)(-\varepsilon^{\alpha\beta\gamma} d_\beta^\dagger \bar{\sigma}_M \bar{u}_\gamma) + (-\varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} \bar{\sigma}^M d^\beta)(\bar{e}^\dagger \bar{\sigma}_M u^\alpha) + h.c.$$

and

$$J^{+4/3} J^{-4/3} = (\bar{d}_\alpha \sigma^M e^\dagger)(+\varepsilon^{\alpha\beta\gamma} u_\beta^\dagger \bar{\sigma}_M \bar{u}_\gamma) + (+\varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} \bar{\sigma}^M u^\beta)(\bar{e}^\dagger \bar{\sigma}_M d^\alpha) + h.c.$$

248

These can be simplified using the identities

$$(\sigma^M)_{m\dot{m}}(\bar{\sigma}_M)^{\dot{n}n} = 2\,\delta^n_m\delta^{\dot{n}}_{\dot{m}} \quad \text{and} \quad (\bar{\sigma}^M)^{\dot{m}m}(\bar{\sigma}_M)^{\dot{n}n} = 2\,\varepsilon^{mn}\varepsilon^{\dot{m}\dot{n}}$$

where we use the metric $\eta = (+,-,-,-)$ and the convention $\varepsilon^{12} = \varepsilon^{\dot{1}\dot{2}} = +1$. Using these, we have $(\xi\sigma^M\nu^\dagger)(\psi^\dagger\bar{\sigma}_M\chi) = -2(\xi\chi)(\psi^\dagger\nu^\dagger)$ and $(\xi^\dagger\bar{\sigma}^M\nu)(\psi^\dagger\bar{\sigma}_M\chi) = +2(\xi^\dagger\psi^\dagger)(\nu\chi)$ for any four Weyl spinors ξ, ν, ψ and χ. (Relations like these are called Fierz identities.) So the above products of currents simplify to

$$J^{+1/3}J^{-1/3} = +2\,\varepsilon^{\alpha\beta\gamma}(\bar{d}_\alpha\bar{u}_\gamma)(d^\dagger_\beta\nu^\dagger) - 2\,\varepsilon_{\alpha\beta\gamma}(\bar{u}^{\dagger\gamma}\bar{e}^\dagger)(d^\beta u^\alpha)$$

$$J^{+4/3}J^{-4/3} = -2\,\varepsilon^{\alpha\beta\gamma}(\bar{d}_\alpha\bar{u}_\gamma)(u^\dagger_\beta e^\dagger) + 2\,\varepsilon_{\alpha\beta\gamma}(\bar{u}^{\dagger\gamma}\bar{e}^\dagger)(u^\beta d^\alpha)\ .$$

Adding these gives the low-energy effective Lagrangian

$$\mathcal{L}^{\Delta B\neq 0}_{\text{eff}} = \frac{2g^2}{M_X^2}\left[\varepsilon^{\alpha\beta\gamma}(\bar{d}_\alpha\bar{u}_\gamma)(d^\dagger_\beta\nu^\dagger - u^\dagger_\beta e^\dagger) - 2\,\varepsilon_{\alpha\beta\gamma}(\bar{u}^{\dagger\gamma}\bar{e}^\dagger)(d^\beta u^\alpha) + h.c.\right]\ .$$

We would now like to interpret this Lagrangian in terms of the low-energy hadron fields. As emphasized on p. 342, any two Lagrangians that exhibit identical symmetries must describe the same low-energy physics. This leads to the replacements[29]

$$\varepsilon_{\alpha\beta\gamma}(\bar{d}^{\dagger\alpha}\bar{u}^{\dagger\beta})(eu^\gamma) \to \mu^3 e\left[p + \frac{i}{f}\left(\pi^+ n + \tfrac{1}{\sqrt{2}}\pi^0 p\right) + O\left(\frac{\pi^2}{f^2}\right)\right]$$

$$\varepsilon_{\alpha\beta\gamma}(\bar{u}^{\dagger\gamma}\bar{e}^\dagger)(d^\alpha u^\beta) \to \mu^3 \bar{e}^\dagger\left[\bar{p}^\dagger - \frac{i}{f}\left(\pi^+\bar{n}^\dagger + \tfrac{1}{\sqrt{2}}\pi^0\bar{p}^\dagger\right) + O\left(\frac{\pi^2}{f^2}\right)\right]$$

where f is the pion decay constant, μ is a strong-interaction parameter with dimensions of mass, π^+ is the charged pion, π^0 is the neutral pion, $(p,\bar{p}^\dagger)$ is the proton and $(n,\bar{n}^\dagger)$ is the neutron, and we drop terms of order π^2/f^2.

Therefore, the part of the baryon-number violating effective Lagrangian relevant for proton decay is

$$\mathcal{L}^{\Delta B\neq 0}_{\text{eff}} = \lambda\left\{e\left[p + \frac{i}{f}\left(\pi^+ n + \tfrac{1}{\sqrt{2}}\pi^0 p\right)\right] + 2\bar{e}^\dagger\left[\bar{p}^\dagger - \frac{i}{f}\left(\pi^+\bar{n}^\dagger + \tfrac{1}{\sqrt{2}}\pi^0\bar{p}^\dagger\right)\right] + h.c.\right\}$$

where we have defined the coupling

$$\lambda \equiv 2g^2\frac{\mu^3}{M_X^2}$$

which has dimensions of mass.

Since all calculations in the text are done with 4-component Dirac spinors instead of 2-component spinors, let us now rewrite $\mathcal{L}^{\Delta B\neq 0}_{\text{eff}}$ in terms of Dirac spinors. Define

$$E^c \equiv \begin{pmatrix} \bar{e} \\ e^\dagger \end{pmatrix}\ ,\quad \mathcal{P} \equiv \begin{pmatrix} p \\ \bar{p}^\dagger \end{pmatrix}\ ,\quad \mathcal{N} \equiv \begin{pmatrix} n \\ \bar{n}^\dagger \end{pmatrix}$$
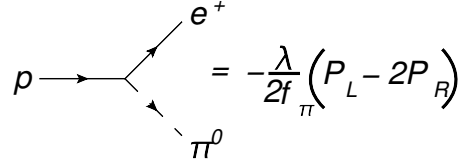
---

[29]See the addendum to this chapter of solutions.

where E^c is the Dirac field for the positron, or in other words the charge-conjugate of the Dirac field $E \equiv \begin{pmatrix} e \\ \bar{e}^\dagger \end{pmatrix}$ for the electron. Since $\overline{E^c}\mathbb{P}_L = ep$ and $\overline{E^c}\mathbb{P}_R = \bar{e}^\dagger \bar{p}^\dagger$, we have:

$$\mathcal{L}_{\text{eff}}^{\Delta B \neq 0} = \lambda \,\overline{E^c}\left[(\mathbb{P}_L + 2\mathbb{P}_R) + \frac{i}{2f_\pi}\pi^0(\mathbb{P}_L - 2\mathbb{P}_R)\right]\mathcal{P} + h.c.$$

$$+ \lambda \,\frac{i}{\sqrt{2}\,f_\pi}\pi^+\overline{E^c}(\mathbb{P}_L - 2\mathbb{P}_R)\mathcal{N} + h.c.$$

where $\mathbb{P}_L \equiv \frac{1}{2}(I - \gamma^5)$ and $\mathbb{P}_R \equiv \frac{1}{2}(I + \gamma^5)$, and we have defined a modified pion decay constant $f_\pi \equiv f/\sqrt{2}$.

We will first compute the rate for $p \to \pi^0 e^+$. We see from the above Lagrangian that one contribution to the amplitude arises from the cubic $\pi^0 e^+ p$ vertex
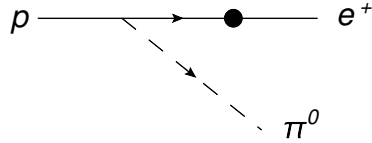

$$= -\frac{\lambda}{2f_\pi}\left(P_L - 2P_R\right)$$

There is also a contribution from the pion-nucleon interactions derived from the chiral Lagrangian:

$$\mathcal{L}_{\pi^0 pn} = \frac{g_A}{2f_\pi}\partial_\mu \pi^0 \left(\overline{\mathcal{P}}\gamma^\mu \gamma^5 \mathcal{P} - \overline{\mathcal{N}}\gamma^\mu \gamma^5 \mathcal{N}\right)$$

where g_A is the axial vector coupling. The contributing diagram to the proton decay amplitude is

where the solid dot indicates the $i\lambda\overline{E^c}(\mathbb{P}_L + 2\mathbb{P}_R)\mathcal{P}$ vertex from $\mathcal{L}_{\text{eff}}^{\Delta B \neq 0}$. Let k_0, k_1, k_2 be the momenta of the proton, positron, and pion respectively. Then the second diagram contributes:

$$\bar{u}_1\left[i\lambda\left(\mathbb{P}_L + 2\mathbb{P}_R\right)\right]\left[iS_p(k_1)\right]\left[-\frac{g_A}{2f_\pi}\slashed{k}_2\gamma^5\right]u_0$$

$$= +\frac{\lambda g_A}{2f_\pi}\,\bar{u}_1(\mathbb{P}_L + 2\mathbb{P}_R)\frac{\slashed{k}_1 + m_p}{k_1^2 - m_p^2}\,\slashed{k}_2\gamma^5 u_0 \ .$$

Using $\gamma^5\gamma^\mu$, we can move the $\slashed{k}_1$ to the left of $\mathbb{P}_L + 2\mathbb{P}_R$, picking up some minus signs. We will neglect the electron mass $m_e \ll m_p$, and so $\bar{u}_1 \slashed{k}_1 = \bar{u}_1 m_e \approx 0$. In the denominator, we also have $k_1^2 = m_e^2 \approx 0$.

Furthermore, we have $k_2 = k_0 - k_1$, so that again we can move $\slashed{k}_1$ to the left and get

250

zero when acted on by $\bar{u}_1$. Finally, using $\mathbb{P}_L\gamma^5 = -\mathbb{P}_L$ and $\mathbb{P}_R\gamma^5 = +\mathbb{P}_R$, we arrive at two powers of m_p in the numerator to cancel out the $1/m_p^2$ from the propagator and obtain:

$$- \frac{\lambda g_A}{2f_\pi}\bar{u}_1(\mathbb{P}_L - 2\mathbb{P}_R)u_0$$

for the total contribution to the amplitude from the second diagram.

This is exactly the same as the contribution from the $\pi^0 e^+ p$ vertex, up to a factor of g_A. The amplitude is

$$i\mathcal{M} = -(1 + g_A)\frac{\lambda}{2f_\pi}\bar{u}_1(\mathbb{P}_L - 2\mathbb{P}_R)u_0 \ .$$

Taking the magnitude squared, averaging over the proton spins and summing over the positron spins using the usual manipulations gives

$$\tfrac{1}{2}\sum_{s_0,s_1}|\mathcal{M}|^2 = \frac{5}{32f_\pi^2}\lambda^2(1 + g_A)^2\frac{m_p^2 - m_\pi^2}{m_p m_e}$$

where we have set $m_e \approx 0$ in the numerator but kept the mass of the neutral pion. The decay rate is (p. 141)

$$\Gamma = \int \frac{d^3k_1}{(2\pi)^3(\omega_1/m_e)}\frac{d^3k_2}{(2\pi)^3 2\omega_2}(2\pi)^4\delta^4(k_0 - k_1 - k_2)\ \tfrac{1}{2}\sum_{s_0,s_1}|\mathcal{M}|^2$$

where $\omega_i = \sqrt{|\vec{p}_i|^2 + m_i^2}$. The amplitude is a constant, and the 2-body phase space integral is given by equation (40) on p. 142 suitably adjusted to account for the fermion factor ω_1/m_e:

$$\int \frac{d^3k_1}{(\omega_1/m_e)}\frac{d^3k_2}{2\omega_2}\delta^4(k_0 - k_1 - k_2) \approx \left(\frac{m_p^2 - m_\pi^2}{m_p^2}\right)\pi m_e$$

where again we have taken $m_e \approx 0$ wherever possible. The decay rate is therefore

$$\Gamma(p \to e^+\pi^0) = \frac{5}{128\pi}\frac{\lambda^2(1 + g_A)^2}{f_\pi^2}\frac{(m_p^2 - m_\pi^2)^2}{m_p^3} \ .$$

Actually this is still not quite complete, since the coefficients of the hadronic operators in the effective Lagrangian are subject to renormalization group corrections. See F. Wilczek and A. Zee, "Operator analysis of nucleon decay," Phys. Rev. Lett. Vol. 43 No. 21, 19 Nov 1979.

Using the effective Lagrangian, the other decay rates $p \to \bar{\nu}\pi^+$, $n \to e^+\pi^-$ and $n \to \bar{\nu}\pi^0$ may be computed in exactly the same way. For general relations between the rates, see problems VIII.3.3 and VIII.3.4.

3. Show that $SU(5)$ conserves the combination $B - L$. For a challenge, invent a grand unified theory that violates $B - L$.

Solution:

We follow F. Wilczek and A. Zee, "Conservation or Violation of $B - L$ in Proton Decay," Phys. Lett. Vol. 88B, No. 3,4 17 Dec 1979.

The Higgs field $\varphi^\mu \sim 5$ of $SU(5)$ couples to the fermion fields $\psi_\mu \sim \bar{5}$ and $\psi^{\mu\nu} \sim 5 \otimes_A 5 = 10$ through the Yukawa interactions given in equations[30] (11) and (12) on p. 417:

$$\mathcal{L} = -f_1\, \psi_\mu C \psi^{\mu\nu} \varphi_\nu^\dagger - f_2\, \psi^{\mu\nu} C \psi^{\lambda\rho} \varphi^\sigma \varepsilon_{\mu\nu\lambda\rho\sigma}$$

where f_1 and f_2 are coupling constants. These terms (and the rest of the Lagrangian) are invariant under an accidental $U(1)$ symmetry. Let us call this symmetry $U(1)_X$ with generator X. Then to leave the Yukawa terms invariant, we require $X(\bar{5}) + X(10) - X(\varphi) = 0$ and $2X(10) + X(\varphi) = 0$, where $X(\bar{5})$ is the $U(1)_X$ charge of the field $\psi_\mu \sim \bar{5}$, and so forth. So $X(\varphi) = -2X(10)$ and thus $X(\bar{5}) = -3X(10)$. Thus choosing the normalization convention $X(10) = 1$, the Lagrangian is invariant under $U(1)_X$ if the fields have charges $X(\psi_\mu) = -3$, $X(\psi^{\mu\nu}) = +1$, $X(\varphi^\mu) = -2$.

After φ^μ acquires a vacuum expectation value, the $U(1)_X$ symmetry gets broken, but a subgroup remains unbroken. Recall that in terms of low-energy fields, the $SU(5)$ fermions are $(\psi_{\bar{5}})_\mu \equiv \begin{pmatrix} \bar{d}_\alpha \\ \ell_i \end{pmatrix}$ for the $\bar{5}$, and for the 10: $(\psi_{10})^{\alpha\beta} = \varepsilon^{\alpha\beta\gamma}\bar{u}_\gamma$, $(\psi_{10})_i^\alpha = q_i^\alpha = \begin{pmatrix} u^\alpha \\ d^\alpha \end{pmatrix}$ and $(\psi_{10})^{ij} = \varepsilon^{ij}\bar{e}$. Again recalling problem VII.5.1, we have $(\psi_{10})^{\alpha i} = \varepsilon^{ij}(\psi_{10})_j^\alpha$ so that $(\psi_{10})^{\alpha 4} = \varepsilon^{45}d^\alpha = +d^\alpha$ and $(\psi_{10})^{\alpha 5} = \varepsilon^{54}u^\alpha = -u^\alpha$.

We can thereby evaluate the generator $\frac{1}{2}Y$ of hypercharge on the ψ_μ and $\psi^{\mu\nu}$ to find that the combination $X + 4(Y/2)$ generates a symmetry even after spontaneous symmetry breaking of $SU(5)$. The other conserved generators are Q, the generator of electric charge, and $\{T^a\}_{a=1}^8$, the generators of color $SU(3)$. By explicitly evaluating the generator $X + 4(Y/2)$ on the components of ψ_μ and $\psi^{\mu\nu}$, we will find that $X + 4(Y/2)$ is a multiple of $B - L$:

$$[X + 4(\tfrac{1}{2}Y)]\bar{d} = [-3 + 4(+\tfrac{1}{3})]\bar{d} = 5(-\tfrac{1}{3})\bar{d} \ (\leftarrow B = -\tfrac{1}{3}, \ L = 0 \ \checkmark)$$
$$[X + 4(\tfrac{1}{2}Y)]\ell = [-3 + 4(-\tfrac{1}{2})]\ell = 5(-1)\ell \ (\leftarrow B = 0, \ L = +1 \ \checkmark)$$
$$[X + 4(\tfrac{1}{2}Y)]\bar{u} = [+1 + 4(-\tfrac{2}{3})]\bar{u} = 5(-\tfrac{1}{3})\bar{u} \ (\leftarrow B = -\tfrac{1}{3}, \ L = 0 \ \checkmark)$$
$$[X + 4(\tfrac{1}{2}Y)]q = [+1 + 4(+\tfrac{1}{6})]q = 5(+\tfrac{1}{3})q \ (\leftarrow B = +\tfrac{1}{3}, \ L = 0 \ \checkmark)$$
$$[X + 4(\tfrac{1}{2}Y)]\bar{e} = [+1 + 4(+1)]\bar{e} = 5(+1)\bar{e} \ (\leftarrow B = 0, \ L = -1 \ \checkmark)$$

Therefore, $B - L = \frac{1}{5}[X + 4(\tfrac{1}{2}Y)]$.

A deeper understanding of this seemingly accidental symmetry can be obtained by embedding the $SU(5)$ theory into $SO(10)$, as explained in the next chapter. One finds that $B - L$

---

[30]We write $\varphi_\mu^\dagger \equiv (\varphi^\mu)^\dagger \sim \bar{5}$ to emphasize the complex conjugation.

is a generator of the $SO(10)$ theory due to the presence of the right-handed gauge-singlet neutrino which can be assigned a lepton number of -1.

One way to violate $B - L$ in the $SU(5)$ theory is to add another Higgs field, H, transforming as $H^{\mu\nu} \sim 5 \otimes_A 5 = 10$. The field H couples to fermions through the Yukawa interaction

$$\mathcal{L} = f_{IJ} \psi_I^\mu C \psi_J^\nu (H^\dagger)_{\mu\nu}$$

where the repeated indices I, J label the different families[31]. At this stage H could simply be assigned a charge under X such that $B - L$ remains conserved, but in the absence of further restrictions there is a cubic scalar interaction:

$$\mathcal{L} = \mu \, H^{\mu\nu} H^{\rho\sigma} \varphi^\lambda \varepsilon_{\mu\nu\rho\sigma\lambda}$$

where μ is a coupling with dimensions of mass. The clash between these two terms violates $B - L$ by two units and thus implies $(B - L)$-violating processes such as $n \to \mu^- K^+$ and $p \to \mu^- K^+ \pi^+$.

Addendum: Chiral Lagrangian for $SU(3) \otimes SU(3)$

Here we review the nonlinear sigma model used to parameterize low-energy QCD, also known as the chiral Lagrangian. For more details, see M. Claudson and M. Wise, "Chiral Lagrangian for deep mine physics," Nucl. Phys. B195 (1982) 297-307 and O. Kaymakcalan, L. Chong-Huah and K. C. Wali, "Chiral Lagrangian for proton decay," Phys. Rev. D, Vol. 29 No. 9, 1 May 1984.

The Lagrangian for three generations of massless quarks is

$$\mathcal{L} = i \sum_{i=1}^3 \left(q_i^\dagger \bar{\slashed{D}} q_i + \bar{q}_i^\dagger \bar{\slashed{D}} \bar{q}_i \right)$$

where $q = (u, d, s)$ and $\bar{q} = (\bar{u}, \bar{d}, \bar{s})$ are the two-component spinors for the up, down and strange quarks, and $\bar{\slashed{D}} \equiv \bar{\sigma}^\mu D_\mu$ is the gauge-covariant derivative for each quark field.

This Lagrangian exhibits the global symmetry $SU(3)_L \otimes SU(3)_R$, under which the quarks transform as $q \sim (3, 1)$ and $q \sim (1, \bar{3})$. We will use the indices $A = 1, 2, 3$ and $A' = 1, 2, 3$ to label the 3-dimensional representations of $SU(3)_L$ and $SU(3)_R$ respectively. We also choose upper indices for the fundamental, 3, and lower indices for the anti-fundamental, $\bar{3}$. There is also the global $U(1)_V$ symmetry $(q, \bar{q}^\dagger) \to e^{-i\theta}(q, \bar{q}^\dagger)$. The axial $U(1)_A : (q, \bar{q}) \to e^{-i\theta}(q, \bar{q})$ is anomalous.

---

[31] Since $H^{\mu\nu} = -H^{\nu\mu}$, the coupling f must be antisymmetric: $f_{IJ} = -f_{JI}$. Thus this interaction couples one generation of fermions to another. This is reminiscent of the interaction $f_{ab} h^+ \varepsilon^{ij} \ell_{ai} \ell_{bj}$ present in the Zee model of neutrino masses. See problem VIII.3.2.

As discussed in the text, the QCD vacuum is supposed to break the global symmetry $SU(3)_L \otimes SU(3)_R$ down to the diagonal subgroup $SU(3)_V$ through the formation of a chiral condensate:

$$\langle q^A \bar{q}_{A'} \rangle = -v^3 \delta^A_{\ A'}$$

where v is a positive parameter with dimensions of mass, and the minus sign arises from the fact that fermions contribute negative energy to the vacuum.

On general principles, breaking a global symmetry group G down to a subgroup H results in massless excitations that parameterize the coset space G/H. In other words, breaking $SU(3)_L \otimes SU(3)_R \to SU(3)_V$ results in Nambu-Goldstone bosons that parameterize the "axial" $SU(3)$ coset space

$$\frac{SU(3)_L \otimes SU(3)_R}{SU(3)_V} = SU(3)_A \ .$$

The broken group $SU(3)_A$ has $3^2 - 1 = 8$ generators. For each broken generator there is one massless particle. These are the parity-odd spinless mesons:

$$\Pi = \begin{pmatrix} \frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta & K^0 \\ K^- & \bar{K}^0 & -\frac{2}{\sqrt{6}}\eta \end{pmatrix} \ .$$

Each field parameterizes a fluctuation along a particular direction in $SU(3)_A$, generated by $T^a = \frac{1}{2}\lambda^a$, where λ^a are the eight Gell-Mann matrices:

$$\lambda^i = \begin{pmatrix} \sigma^i & 0 \\ 0 & 0 \end{pmatrix} \ , \quad \lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \ , \quad \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}$$

$$\lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \ , \quad \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \ , \quad \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \ .$$

The σ^i are the Pauli matrices:

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ .$$

We may map a spacetime point x^μ to any point on the group manifold of $SU(3)_A$ by treating the components of the meson octet as fields and exponentiating them to form a 3-by-3 special unitary spacetime-dependent matrix:

$$\Sigma(x) \equiv e^{i2\Pi(x)/f}$$

where f is the pion decay constant, $f \approx 139$ MeV. Under a general $SU(3)_L \otimes SU(3)_R$ transformation, the matrix Σ transforms as $\Sigma \to L\Sigma R^\dagger$, where L is an element of $SU(3)_L$ and R is an element of $SU(3)_R$. The diagonal subgroup $SU(3)_V$ is given by transformations for

which $L = R$.

It will prove convenient to define the "square root" of the matrix Σ as $\xi(x) \equiv e^{i\Pi(x)/f}$. Under a general $SU(3)_L \otimes SU(3)_R$ transformation, the matrix ξ is defined to transform as

$$\xi \to L\xi U^\dagger = U\xi R^\dagger$$

where U is a 3-by-3 unitary matrix that is defined by the above transformation law and thereby depends nonlinearly on the matrices L, R and Π. This ensures that $\Sigma = \xi^2$ transforms as stated.

After seeing how the mesons arise from chiral symmetry breaking, we need the spectrum of baryons at low energy.[32] The quark fields transform under $SU(3)_L \otimes SU(3)_R$ as

$$q \equiv \begin{pmatrix} u \\ d \\ s \end{pmatrix} \sim (3,1) \quad \text{and} \quad \bar{q} \equiv \begin{pmatrix} \bar{u} \\ \bar{d} \\ \bar{s} \end{pmatrix} \sim (1,\bar{3}) \ .$$

Therefore $\bar{q}^\dagger$ transforms as $(1,3)$, and

$$\bar{q}^\dagger \bar{q}^\dagger \sim (1,3) \otimes (1,3) = (1,\bar{3}_A) \oplus (1,6_S) \ .$$

A baryon is a color-singlet bound state of three quarks, so the color indices of the above product of quarks are to be contracted with the 3-index antisymmetric tensor of $SU(3)_c$. The Lorentz-singlet part[33] of $\bar{q}^\dagger \bar{q}^\dagger$ forms a color anti-triplet

$$Q_\alpha \equiv \tfrac{1}{2}\varepsilon_{\alpha\beta\gamma}\bar{q}^{\dagger\beta}\bar{q}^{\dagger\gamma} = \begin{pmatrix} \bar{d}^{\dagger\beta}\bar{s}^{\dagger\gamma} \\ \bar{s}^{\dagger\beta}\bar{u}^{\dagger\gamma} \\ \bar{u}^{\dagger\beta}\bar{d}^{\dagger\gamma} \end{pmatrix}$$

which transforms as $(1,\bar{3})$ under global $SU(3)_L \otimes SU(3)_R$ transformations. (Here an upper $\alpha = 1,2,3$ denotes the fundamental of color and a lower index α denotes the anti-fundamental.)

Thus the color-singlet composite field

$$\vec{q}^{\,\alpha}(\vec{Q}_\alpha)^T = \varepsilon_{\alpha\beta\gamma} \begin{pmatrix} u^\alpha(\bar{d}^{\dagger\beta}\bar{s}^{\dagger\gamma}) & u^\alpha(\bar{s}^{\dagger\beta}\bar{u}^{\dagger\gamma}) & u^\alpha(\bar{u}^{\dagger\beta}\bar{d}^{\dagger\gamma}) \\ d^\alpha(\bar{d}^{\dagger\beta}\bar{s}^{\dagger\gamma}) & d^\alpha(\bar{s}^{\dagger\beta}\bar{u}^{\dagger\gamma}) & d^\alpha(\bar{u}^{\dagger\beta}\bar{d}^{\dagger\gamma}) \\ s^\alpha(\bar{d}^{\dagger\beta}\bar{s}^{\dagger\gamma}) & s^\alpha(\bar{s}^{\dagger\beta}\bar{u}^{\dagger\gamma}) & s^\alpha(\bar{u}^{\dagger\beta}\bar{d}^{\dagger\gamma}) \end{pmatrix}$$

transforms as $(3,\bar{3})$ under global $SU(3)_L \otimes SU(3)_R$ transformations. From the electric charges $(+\tfrac{2}{3}, -\tfrac{1}{3}, -\tfrac{1}{3})$ of (up, down, strange), the electric charges of the components of the above composite field are

$$\begin{pmatrix} 0 & +1 & +1 \\ -1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \ .$$

---

[32] We thank Tim Tait and Mark Srednicki for helpful discussions on this.

[33] We are considering the 2-component Lorentz-spinor indices contracted as $\bar{u}^\dagger \bar{d}^\dagger \equiv \bar{u}_{\dot{a}}^\dagger \bar{d}^{\dagger\dot{a}} \equiv \bar{u}_{\dot{a}}^\dagger \varepsilon^{\dot{a}\dot{c}} \bar{d}_{\dot{c}}^\dagger$, so that $\bar{u}^\dagger \bar{d}^\dagger = \bar{d}^\dagger \bar{u}^\dagger$ for Grassmann-valued fermion fields $\bar{u}_{\dot{a}}^\dagger$ and $\bar{d}_{\dot{a}}^\dagger$.

At low energy, the QCD vacuum breaks $SU(3)_L \otimes SU(3)_R \to SU(3)_V$, under which $(3, \bar{3}) \to 3 \otimes \bar{3} = 1 \otimes 8$. We therefore find an octet of baryons[34]

$$B = \begin{pmatrix} \frac{1}{\sqrt{2}}\Sigma^0 + \frac{1}{\sqrt{6}}\Lambda^0 & \Sigma^+ & p \\ \Sigma^- & -\frac{1}{\sqrt{2}}\Sigma^0 + \frac{1}{\sqrt{6}}\Lambda^0 & n \\ \Xi^- & \Xi^0 & -\frac{2}{\sqrt{6}}\Lambda^0 \end{pmatrix}$$

given by the replacement $\vec{q}^{\,\alpha}(\vec{Q}_\alpha)^T - \frac{1}{3}I\mathrm{tr}[\vec{q}^{\,\alpha}(\vec{Q}_\alpha)^T] \to \mu^3 B$, where μ is a parameter with dimensions of mass and I is the 3-by-3 identity matrix. For example, the proton and neutron are $p \sim u(\bar{u}^\dagger \bar{d}^\dagger)$ and $n \sim d(\bar{u}^\dagger \bar{d}^\dagger)$, as they should be.

There is also the octet of anti-baryons, $\bar{B}$, which can be derived analogously. The baryon Dirac field is given by $\mathcal{B} = \begin{pmatrix} B \\ \bar{B}^\dagger \end{pmatrix}$.

Under a general $SU(3)_L \otimes SU(3)_R$ transformation, the baryon octet transforms as $B \to UBU^\dagger$, where again the matrix U is defined by the transformation law $\xi \to L\xi U^\dagger = U\xi R^\dagger$. Therefore, we have $\xi B \xi \to L \xi B \xi R^\dagger$ so that $\xi B \xi$ transforms as $\Sigma = \xi^2$. Similarly, $\xi^\dagger B \xi^\dagger$ transforms as $\Sigma^\dagger$.

The next step is to write all possible interaction terms between baryons and mesons that respect $SU(3)_L \otimes SU(3)_R$ symmetry and parity. The result is

$$\begin{aligned} \mathcal{L}_{\pi B} = &\tfrac{1}{8}f^2 \mathrm{tr}(\partial_\mu \Sigma \partial^\mu \Sigma^\dagger) + \mathrm{tr}[\bar{\mathcal{B}}(i\slashed{\partial} - m)\mathcal{B}] \\ &+ \tfrac{1}{2}i\,\mathrm{tr}\left[\bar{\mathcal{B}}\gamma^\mu(\xi\partial_\mu\xi^\dagger + \xi^\dagger\partial_\mu\xi)\mathcal{B} + \mathcal{B}(\partial_\mu\xi\,\xi^\dagger + \partial_\mu\xi^\dagger\,\xi)\right] \\ &- \tfrac{1}{2}i(D - F)\,\mathrm{tr}\left[\bar{\mathcal{B}}\gamma^\mu\gamma^5\mathcal{B}(\partial_\mu\xi\,\xi^\dagger - \partial_\mu\xi^\dagger\,\xi)\right] \\ &+ \tfrac{1}{2}i(D + F)\,\mathrm{tr}\left[\bar{\mathcal{B}}\gamma^\mu\gamma^5(\xi\partial_\mu\xi^\dagger - \xi^\dagger\partial_\mu\xi)\mathcal{B}\right]\ . \end{aligned}$$

The factor of $\frac{1}{8}f^2$ is so that the meson fields that comprise $\Sigma = e^{i2\Pi/f}$ have canonically normalized kinetic terms, e.g. $\frac{1}{2}\partial_\mu\pi^0\partial^\mu\pi^0$ and $\partial_\mu\pi^+\partial^\mu\pi^-$. The parameters $F = 0.44$ and $D = 0.81$ are measured from semileptonic baryon decays.

Expanding $\mathcal{L}_{\pi B}$ using $\xi = I + i\Pi/f + ...$ and the components of the baryon octet, we find interactions between the protons, neutrons and pions:

$$\begin{aligned} \mathcal{L}_{\pi B} = &\tfrac{1}{2}\partial_\mu\pi^0\partial^\mu\pi^0 + \partial_\mu\pi^+\partial^\mu\pi^- + \overline{\mathcal{P}}(i\slashed{\partial} - m)\mathcal{P} + \overline{\mathcal{N}}(i\slashed{\partial} - m)\mathcal{N} \\ &+ \left(\frac{D+F}{f}\right)\left[\tfrac{1}{\sqrt{2}}\partial_\mu\pi^0\left(\overline{\mathcal{P}}\gamma^\mu\gamma^5\mathcal{P} - \overline{\mathcal{N}}\gamma^\mu\gamma^5\mathcal{N}\right) + \left(\partial_\mu\pi^+\overline{\mathcal{P}}\gamma^\mu\gamma^5\mathcal{N} + h.c.\right)\right] + ... \end{aligned}$$

where the "..." stand for terms involving strange mesons and baryons, as well as terms of higher order in $\partial_\mu\pi/f$. It is often customary to define the axial vector coupling $g_A = D + F$

---

[34] We also find an $SU(3)_V$-singlet, electrically-neutral baryon $\frac{1}{\sqrt{3}}\mathrm{tr}(\vec{q}\,\vec{Q}^T) = \frac{1}{\sqrt{3}}[u(\bar{d}^\dagger\bar{s}^\dagger) + d(\bar{s}^\dagger\bar{u}^\dagger) + s(\bar{u}^\dagger\bar{d}^\dagger)]$, where we have suppressed the color indices.

and a modified pion decay constant $f_\pi = f/\sqrt{2}$, so that the coupling of the neutral pion to the proton and neutron can be written as

$$\mathcal{L}_{\pi^0 pn} = \frac{g_A}{2f_\pi} \partial_\mu \pi^0 \left(\overline{\mathcal{P}} \gamma^\mu \gamma^5 \mathcal{P} - \overline{\mathcal{N}} \gamma^\mu \gamma^5 \mathcal{N} \right)$$
$$= \frac{g_A}{f_\pi} \partial_\mu \pi^0 \, \overline{\Psi} \gamma^\mu \gamma^5 I_3 \Psi \ ,$$

where $\Psi \equiv \begin{pmatrix} \mathcal{P} \\ \mathcal{N} \end{pmatrix}$ is the nucleon doublet, and $I_3 \equiv \begin{pmatrix} +\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$ is the third generator of isospin.

Introducing nonzero quark masses results in nonzero masses for the meson octet and splits the degeneracy among the masses in the baryon octet. For further details, see the references.

VII.7 $SO(10)$ **Unification**

1. Work out the Clifford algebra in d-dimensional space for d odd.

Solution:

The Clifford algebra for $d = 2n + 1$ is that of $d = 2n$ with the addition of the γ^{FIVE} from $d = 2n$. For example, consider $d = 3$. The gamma matrices for $d = 2$ are just $\gamma^1 = \sigma^1$ and $\gamma^2 = \sigma^2$ as explained in the text. The chirality matrix $\gamma^{\mathrm{FIVE}} \equiv -i\gamma^1\gamma^2 = -i\sigma^1\sigma^2 = \sigma^3$ anticommutes with γ^1 and γ^2, and it squares to 1. Therefore the Clifford algebra for $d = 3$ is

$$\{\gamma^i, \gamma^j\} = 2\delta^{ij} \ , \quad \gamma^i = \sigma^i$$

which you already encountered way back in problem II.1.12, where you discovered that the Dirac mass term in $(2 + 1)$-dimensional spacetime violates parity and time reversal. In general, the γ^{FIVE} for $d = 2n$ anticommutes with the $\{\gamma^i\}_{i=1}^{2n}$ from $d = 2n$ (since $2n$ is always even) and squares to 1, and therefore will form a perfectly good $(2n + 1)^{\mathrm{th}}$ gamma matrix $\gamma^{2n+1} \equiv \gamma^{\mathrm{FIVE}} \equiv (-i)^n \gamma^1 \gamma^2 ... \gamma^{2n}$ for $d = 2n + 1$. The point is that while the spinor representation of $SO(2n)$ is reducible into two chiral irreducible representations, the spinor representation of $SO(2n + 1)$ is not reducible.

2. Work out the Clifford algebra in d-dimensional Minkowski space.

Solution:

The defining equation of the Clifford algebra in d-dimensional Minkowski space is

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \ , \quad \eta^{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1, ..., -1)$$

where $\mu, \nu = 0, 1, 2, ..., d-1$. Let $i, j = 1, 2, ..., d-1$ denote purely spatial indices. Then the above can be written as

$$(\gamma^0)^2 = +1, \ (\gamma^i)^2 = -1, \ \{\gamma^i, \gamma^j\} = -2\delta^{ij}$$

All we have to do is to take the Clifford algebra for d-dimensional Euclidean space and throw in some factors of i to generate the appropriate minus signs in the metric. Immediately we find

$$\gamma^0_{SO(d-1,1)} = \gamma^0_{SO(d)} , \quad \gamma^j_{SO(d-1,1)} = i\gamma^j_{SO(d)}$$

for $j = 1, ..., d-1$. Note that we are using slightly different notation for the $SO(d)$ gamma matrices from the text. In the chapter, the vector indices run from 1 to d, whereas here our indices run from 0 to $d-1$. If you want to stick to the notation in the book, you can write

$$\gamma^0_{SO(d-1,1)} = \gamma^1_{SO(d)} , \quad \gamma^j_{SO(d-1,1)} = i\gamma^{j+1}_{SO(d)}$$

for $j = 1, ..., d-1$.

3. Show that the Clifford algebra for $d = 4k$ and for $d = 4k+2$ have somewhat different properties. (If you need help with this and the two preceding exercises, look up F. Wilczek and A. Zee, *Phys. Rev.* D25: 553, 1982.)

Solution:

The chirality matrix $\gamma^{\text{FIVE}} = \sigma_3 \otimes ... \otimes \sigma_3$ acts on a chiral spinor $|\varepsilon_1, ..., \varepsilon_n\rangle$ as $\gamma^{\text{FIVE}}|\varepsilon_1, ..., \varepsilon_n\rangle = \left(\prod_{j=1}^n \varepsilon_j\right)|\varepsilon_1, ..., \varepsilon_n\rangle$. The "left-handed" chiral spinor $|\varepsilon_1, ..., \varepsilon_n\rangle_L$ is defined by $\gamma^{\text{FIVE}}|\varepsilon_1, ..., \varepsilon_n\rangle = (-1)|\varepsilon_1, ..., \varepsilon_n\rangle$, and the "right-handed" chiral spinor $|\varepsilon_1, ..., \varepsilon_n\rangle_R$ is defined by $\gamma^{\text{FIVE}}|\varepsilon_1, ..., \varepsilon_n\rangle_R = (+1)|\varepsilon_1, ..., \varepsilon_n\rangle_R$. That is, the sign of $\left(\prod_{j=1}^n \varepsilon_j\right)$ determines whether the spinor is left-handed or right-handed.

Now, the charge conjugation matrix $C = i\sigma_2 \otimes ... \otimes i\sigma_2$ acts on the $SO(2n)$ chiral spinor $|\varepsilon_1, ..., \varepsilon_n\rangle$ as $C|\varepsilon_1, ..., \varepsilon_n\rangle = |(-\varepsilon_1), ..., (-\varepsilon_n)\rangle$, which means that the chirality matrix γ^{FIVE} acts on a charge-conjugated spinor as

$$\gamma^{\text{FIVE}}C|\varepsilon_1, ..., \varepsilon_n\rangle =$$

$$\gamma^{\text{FIVE}}|(-\varepsilon_1), ..., (-\varepsilon_n)\rangle = \left(\prod_{j=1}^n (-\varepsilon_j)\right)|(-\varepsilon_1), ..., (-\varepsilon_n)\rangle = (-1)^n C|\varepsilon_1, ..., \varepsilon_n\rangle .$$

Let's process what this means. Consider the case for which n is even, so that $(-1)^n = +1$. For that case, if $\gamma^{\text{FIVE}}|\varepsilon_1, ..., \varepsilon_n\rangle = +\gamma^{\text{FIVE}}C|\varepsilon_1, ..., \varepsilon_n\rangle$, then $\gamma^{\text{FIVE}}C|\varepsilon_1, ..., \varepsilon_n\rangle = +C|\varepsilon_1, ..., \varepsilon_n\rangle$. If $\gamma^{\text{FIVE}}|\varepsilon_1, ..., \varepsilon_n\rangle = -\gamma^{\text{FIVE}}C|\varepsilon_1, ..., \varepsilon_n\rangle$, then $\gamma^{\text{FIVE}}C|\varepsilon_1, ..., \varepsilon_n\rangle = -C|\varepsilon_1, ..., \varepsilon_n\rangle$. That is, $|\varepsilon_1, ..., \varepsilon_n\rangle$ and $C|\varepsilon_1, ..., \varepsilon_n\rangle$ have the same eigenvalue under γ^{FIVE}; the states $|\varepsilon_1, ..., \varepsilon_n\rangle_L$ and $|\varepsilon_1, ..., \varepsilon_n\rangle_R$ are therefore self-conjugate (that is, not conjugate to each other) if n is even.

Now consider the case for which n is odd, meaning $(-1)^n = -1$. Under that assumption, we have the opposite situation from before. If $\gamma^{\text{FIVE}}|\varepsilon_1, ..., \varepsilon_n\rangle = +\gamma^{\text{FIVE}}C|\varepsilon_1, ..., \varepsilon_n\rangle$, then $\gamma^{\text{FIVE}}C|\varepsilon_1, ..., \varepsilon_n\rangle = -C|\varepsilon_1, ..., \varepsilon_n\rangle$. If $\gamma^{\text{FIVE}}|\varepsilon_1, ..., \varepsilon_n\rangle = -\gamma^{\text{FIVE}}C|\varepsilon_1, ..., \varepsilon_n\rangle$, then we have $\gamma^{\text{FIVE}}C|\varepsilon_1, ..., \varepsilon_n\rangle = +C|\varepsilon_1, ..., \varepsilon_n\rangle$. That is, $|\varepsilon_1, ..., \varepsilon_n\rangle$ and $C|\varepsilon_1, ..., \varepsilon_n\rangle$ have opposite eigenvalues under γ^{FIVE}; the states $|\varepsilon_1, ..., \varepsilon_n\rangle_L$ and $|\varepsilon_1, ..., \varepsilon_n\rangle_R$ are therefore conjugate to each other if n is odd.

Rephrasing somewhat, let $d = 2n$. If n is even, then $n = 2k$ for some integer k. We have shown that the chiral spinors are self-conjugate when n is even, or in other words when $d = 4k = 0$ mod 4. If n is odd, then $n = 2k + 1$ for some integer k. We have shown that the chiral spinors are conjugate to each other when n is odd, or in other words when $d = 4k + 2 = 2$ mod 4.

As an important aside, note that the situation is reversed for $SO(d-1, 1)$, that is for d-dimensional Minkowski spacetime rather than d-dimensional Euclidean spacetime. For $d = 4k$ (that is, $d = 0$ mod 4), the two chiral spinor representations are conjugate to each other. For $d = 2 + 4k$ ($d = 2$ mod 4), each chiral spinor is its own conjugate.

If you desire a second reference for this problem (and other properties of $SO(2n)$ and $SO(2n-1, 1)$ spinors), consult Volume II, Appendix B of J. Polchinski's textbook on string theory.

4. Discuss the Higgs sector of the $SO(10)$. What do you need to give mass to the quarks and leptons?

Solution:

Let $\psi_a \sim 2^{n-1}$ (with $a = 1, ..., 2^{n-1}$) be the $SO(2n)$ spinor that contains a single family of matter fields. (We will specialize to $n = 5$ later.) The $SO(2n)$ invariant tensors at our disposal are the charge conjugation matrix C^{ab} and its inverse C_{ab}, and the gamma matrices $(\gamma^\mu)_a{}^b$, where $\mu = 1, ..., 2n$ denotes the index for the vector ("defining") representation of $SO(2n)$. Note that once we pick the convention for which the spinor ψ has one index down, the rest of the index placements are fixed. For example, an $SO(2n)$ transformation acts as $\psi \to e^{i\omega_{\mu\nu}\sigma^{\mu\nu}}\psi = \psi + i\omega_{\mu\nu}\sigma^{\mu\nu}\psi + O(\omega^2)$, so $\delta\psi_a = i\omega_{\mu\nu}(\sigma^{\mu\nu})_a{}^b\psi_b$. Since $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$, the index placements must be $(\gamma^\mu)_a{}^b$. We can raise and lower indices using C and C^{-1}, which is the whole point of defining such a matrix in the first place. The conventions we use are that indices are raised and lowered by contracting with C or C^{-1} always on the second index: $\psi^a \equiv C^{ab}\psi_b$ and $\psi_a \equiv C_{ab}\psi^b$. This will introduce some minus signs, since in $SO(2n)$ we have the property

$$C^T = (-1)^{n(n+1)/2}C \qquad (A11)$$

For a thorough discussion of this point as well as for other aspects of $SO(2n)$, consult the

paper "Families from Spinors," Phys. Rev. D25: 553, 1982 by F. Wilczek and A. Zee, as referenced in problem VII.7.3. (We label the above equation as (A11) since that is the equation number in the reference.)

To write a mass term for the fermions contained within the spinor ψ_a, we need a term quadratic in ψ that is invariant under $SO(10)$ (and under the Lorentz group). For a single family of fermions, we are restricted to terms of the form

$$\mathcal{L} = -y\,\psi_a C^{ab}(\gamma^{\mu_1}...\gamma^{\mu_K})_b{}^c\psi_c\,\phi_{\mu_1...\mu_K} + h.c.$$

where $\phi_{\mu_1...\mu_K}$ is a Lorentz-scalar field that is completely antisymmetric in its $SO(2n)$ vector indices. (y is just a coupling constant.) We will suppress the Lorentz-spinor indices using the conventions in Appendix E.

From the property $(A11)$, raising and lowering a pair of up-and-down indices introduces the sign $(-1)^{n(n+1)/2}$, meaning: $\psi^a\psi_a = (-1)^{n(n+1)/2}\psi_a\psi^a$, where we remind you that $\psi^a \equiv C^{ab}\psi_b$ and $\psi_b \equiv C_{bc}\psi^c$. This means that raising and lowering *two* pairs of up-and-down indices can always be done without changing a sign regardless of the value of n. In other words, for any matrix $M_a{}^b$, we have $\psi^a M_a{}^b\psi_b = \psi_a M^a{}_b\psi^b$. So the above Lagrangian can equally well be written as

$$\mathcal{L} = -y\,\psi^a C_{ab}(\gamma^{\mu_1}...\gamma^{\mu_K})^b{}_c C^{cd}\psi_d\,\phi_{\mu_1...\mu_K} + h.c.$$

The matrix sandwiched between the ψs is $C^{-1}(\gamma^{\mu_1}...\gamma^{\mu_K})^T C$. We can use the property

$$C^{-1}(\gamma^\mu)^T C = (-1)^n \gamma^\mu \qquad (A10)$$

to simplify this expression:

$$
\begin{aligned}
C^{-1}(\gamma^{\mu_1}...\gamma^{\mu_K})^T C &= C^{-1}(\gamma^{\mu_K})^T...(\gamma^{\mu_1})^T C \\
&= C^{-1}(\gamma^{\mu_K})^T C C^{-1}...C C^{-1}(\gamma^{\mu_1})^T C \\
&= (-1)^{Kn}\gamma^{\mu_K}...\gamma^{\mu_1} \\
&= (-1)^{Kn}(-1)^{K-1}\gamma^{\mu_1}\gamma^{\mu_K}\gamma^{\mu_{K-1}}...\gamma^{\mu_2} \\
&= (-1)^{Kn}(-1)^{K-1}(-1)^{K-2}\gamma^{\mu_1}\gamma^{\mu_2}\gamma^{\mu_K}\gamma^{\mu_{K-1}}...\gamma^{\mu_3} \\
&= ...(\text{keep on anticommuting})... \\
&= (-1)^{Kn}(-1)^{\sum_{j=1}^{K-1} j}\gamma^{\mu_1}...\gamma^{\mu_K} \\
&= (-1)^{Kn}(-1)^{K(K-1)/2}\gamma^{\mu_1}...\gamma^{\mu_K}
\end{aligned}
$$

Therefore:

$$\psi^a C_{ab}(\gamma^{\mu_1}...\gamma^{\mu_K})^b{}_c C^{cd}\psi_d = (-1)^{Kn+K(K-1)/2}\psi^a(\gamma^{\mu_1}...\gamma^{\mu_K})_a{}^d\psi_d$$

Now recall what we said about contracting indices with the second index of the charge conjugation matrix:

$$\psi^a \equiv C^{ab}\psi_b = \psi_b C^{ab} = \psi_b(-1)^{n(n+1)/2}C^{ba}\ .$$

260

Finally, after relabeling some dummy indices, we have the result

$$\psi^a C_{ab}(\gamma^{\mu_1}...\gamma^{\mu_K})^b{}_c C^{cd}\psi_d = (-1)^{Kn+K(K-1)/2+n(n+1)/2}\psi^a C_{ab}(\gamma^{\mu_1}...\gamma^{\mu_K})^b{}_c C^{cd}\psi_d$$

In other words, the mass term for ψ can only be nonzero when the quantity

$$Kn + \frac{K(K-1)}{2} + \frac{n(n+1)}{2} = \frac{(n+K)(n+K+1)}{2} - K$$

is an even number. (This is (A55) in the reference.) Since odd+odd = even and even+even = even while odd+even = odd, this implies that K and $\frac{1}{2}(n+K)(n+K+1)$ are both even or are both odd. Since odd×odd = odd, odd×even = even and even×even = even, we will always have $(n+K)(n+K+1)$ = even. So the question is where $\frac{1}{2}(n+K)$ is even or odd.

$$\frac{n+K}{2} = \text{even} \equiv 2p \implies K = 4p - n$$
$$\frac{n+K}{2} = \text{odd} \equiv 2q+1 \implies K = 4q + 2 - n$$

where p and q are any integers, and of course we must have $K > 0$. The case at hand is $SO(10)$, which means $n = 5$. This implies

$$K = 4p - 5 = 3, 7, 11, ... \quad \left(\frac{5+K}{2} = \text{even}\right)$$
$$K = 4p + 2 - 5 = 1, 5, 9, ... \quad \left(\frac{5+K}{2} = \text{odd}\right)$$

Moreover, recall that $SO(10)$ has the invariant tensor $\varepsilon_{\mu_1...\mu_{10}}$. That means any number of antisymmetrized indices greater than 5 is equivalent to a number less than 5. For example, the definition $\phi_{\mu_1} = \varepsilon_{\mu_1\mu_2...\mu_{10}}\phi_{\mu_2...\mu_9}$ is a statement invariant under $SO(10)$, so $K = 9$ is equivalent to $K = 1$. So the above is really

$$K = 4p - 5 = 3 \quad \left(\frac{5+K}{2} = \text{even}\right)$$
$$K = 4p + 2 - 5 = 1, 5 \quad \left(\frac{5+K}{2} = \text{odd}\right)$$

for $SO(10)$. However, the first of these statements is a contradiction. Recall that we said we must have $\frac{1}{2}(5+K)(6+K)$ and K either both even or both odd. For $K = 3$, we have $\frac{1}{2}(5+3)(6+3) = \frac{1}{2} \times 8 \times 9 = 36$, which is even. Finally we have deduced that only the terms for which $K = 1$ or 5 are allowed in the Lagrangian. (This is the conclusion reached below (A56b) in the reference.) In other words, the masses for one family of fermions comes from the Lagrangian

$$\mathcal{L} = -y\,\psi C\gamma^\mu\psi\,\phi_\mu - y'\,\psi C\gamma^{\mu_1}...\gamma^{\mu_5}\psi\,\phi'_{\mu_1...\mu_5} + h.c.$$

where ϕ and ϕ' are Lorentz-scalar fields, and y and y' are coupling constants. From looking at the indices (and recalling that they are all implicitly antisymmetrized), we see the transformation properties $\phi_\mu \sim 10$ and $\phi'_{\mu_1...\mu_5} \sim 10 \otimes_A 10 \otimes_A 10 \otimes_A 10 \otimes_A 10 = 126$ of $SO(10)$.

Note that for multiple families of fermions, the analysis becomes more complicated. The mass terms have the form

$$\mathcal{L} = -\sum_{F,F'} y_{FF'} \psi_F C \gamma^{\mu_1}...\gamma^{\mu_K} \psi_{F'} \phi_{\mu_1...\mu_K} + h.c.$$

where F and F' run from 1 to the number of fermion families. Lorentz-scalars ϕ satisfying the previous constraints would contribute terms for which $y_{FF'} = y_{F'F}$, and those not satisfying those constraints would contribute terms for which $y_{FF'} = -y_{F'F}$.

5. The group $SO(6)$ has $6(6-1)/2 = 15$ generators. Notice that the group $SU(4)$ also has $4^2 - 1 = 15$ generators. Substantiate your suspicion that $SO(6)$ and $SU(4)$ are isomorphic. Identify some low dimensional representations.

Solution:

The iterative construction of the spinor representation of $SO(2n)$ is given on p. 422. In particular, for $n = 3$ we obtain the 2^3-dimensional spinor representation of $SO(6)$, which is generated by the $\frac{1}{2}(6 \times 5) = 15$ hermitian matrices $\sigma^{\mu\nu} = i\frac{1}{2}[\gamma^\mu, \gamma^\nu]$, with gamma matrices given by:

$$\gamma^{\mu=0,1,2,3} = \gamma^\mu_{(4)} \otimes \sigma_3 = \begin{pmatrix} \gamma^\mu_{(4)} & 0 \\ 0 & -\gamma^\mu_{(4)} \end{pmatrix}$$

$$\gamma^4 = I_{(4)} \otimes \sigma_1 = \begin{pmatrix} 0_{(4)} & I_{(4)} \\ I_{(4)} & 0_{(4)} \end{pmatrix}$$

$$\gamma^5 = I_{(4)} \otimes \sigma_2 = \begin{pmatrix} 0_{(4)} & -iI_{(4)} \\ +iI_{(4)} & 0_{(4)} \end{pmatrix} .$$

This representation is reducible, since we can define the matrix

$$\begin{aligned}
\gamma^{\text{FIVE}} &\equiv -i\gamma^0\gamma^1\gamma^2\gamma^3\gamma^4\gamma^5 \\
&= +i(\gamma^{\text{FIVE}}_{(4)} \otimes \sigma_3^4)(I_{(4)} \otimes \sigma_1\sigma_2) \\
&= +i(\gamma^{\text{FIVE}}_{(4)} \otimes I)(I_{(4)} \otimes i\sigma_3) \\
&= -\gamma^{\text{FIVE}}_{(4)} \otimes \sigma_3 \\
&= \begin{pmatrix} -\gamma^{\text{FIVE}}_{(4)} & 0_{(4)} \\ 0_{(4)} & +\gamma^{\text{FIVE}}_{(4)} \end{pmatrix}
\end{aligned}$$

which commutes with $\sigma^{\mu\nu}$. Thus if Ψ transforms as a 2^3-dimensional spinor, the chiral spinors $\Psi_L \equiv \frac{1}{2}(1 - \gamma^{\text{FIVE}})\Psi$ and $\Psi_R \equiv \frac{1}{2}(1 + \gamma^{\text{FIVE}})\Psi$ have $2^{3-1} = 4$ components and transform irreducibly under $SO(6)$.

Since the chiral spinors are 4-dimensional, the Lorentz generators $\sigma^{\mu\nu}$ acting on them should split up into two sets of fifteen 4-by-4 traceless hermitian matrices. The generators of $SU(4)$

are precisely the set of fifteen 4-by-4 traceless hermitian matrices, so that constructing the irreducible rotation generators acting on the chiral spinor would constitute a proof that $SO(6)$ is locally isomorphic to $SU(4)$.

The rotation generators $\sigma^{\mu\nu}$ take the following forms in terms of 4-dimensional matrices. For $\mu, \nu = 0, 1, 2, 3$ only, we have:

$$
\begin{aligned}
\sigma^{\mu\nu} &= i\tfrac{1}{2}[\gamma^\mu, \gamma^\nu] = i\tfrac{1}{2}[\gamma^\mu_{(4)} \otimes \sigma_3, \gamma^\nu_{(4)} \otimes \sigma_3] \\
&= i\tfrac{1}{2}\left(\gamma^\mu_{(4)}\gamma^\nu_{(4)} \otimes \sigma_3^2 - \gamma^\nu_{(4)}\gamma^\mu_{(4)} \otimes \sigma_3^2\right) \\
&= i\tfrac{1}{2}[\gamma^\mu_{(4)}, \gamma^\nu_{(4)}] \otimes I \\
&= \sigma^{\mu\nu}_{(4)} \otimes I = \begin{pmatrix} \sigma^{\mu\nu}_{(4)} & 0_{(4)} \\ 0_{(4)} & \sigma^{\mu\nu}_{(4)} \end{pmatrix} \quad (\mu, \nu = 0, 1, 2, 3 \text{ only}).
\end{aligned}
$$

For $\mu = 0, 1, 2, 3$, we also have:

$$
\begin{aligned}
\sigma^{\mu 4} &= i\gamma^\mu\gamma^4 = i\left(\gamma^\mu_{(4)} \otimes \sigma_3\right)\left(I_{(4)} \otimes \sigma_1\right) \\
&= i\gamma^\mu_{(4)} \otimes i\sigma_2 = \begin{pmatrix} 0_{(4)} & i\gamma^\mu_{(4)} \\ -i\gamma^\mu_{(4)} & 0_{(4)} \end{pmatrix} \quad (\mu = 0, 1, 2, 3 \text{ only})
\end{aligned}
$$

and

$$
\begin{aligned}
\sigma^{\mu 5} &= i\gamma^\mu\gamma^5 = i\left(\gamma^\mu_{(4)} \otimes \sigma_3\right)\left(I_{(4)} \otimes \sigma_2\right) \\
&= i\gamma^\mu_{(4)} \otimes (-i\sigma_1) = \begin{pmatrix} 0_{(4)} & \gamma^\mu_{(4)} \\ \gamma^\mu_{(4)} & 0_{(4)} \end{pmatrix} \quad (\mu = 0, 1, 2, 3 \text{ only}).
\end{aligned}
$$

Finally, we also have

$$
\begin{aligned}
\sigma^{45} &= i\gamma^4\gamma^5 = i\left(I_{(4)} \otimes \sigma_1\right)\left(I_{(4)} \otimes \sigma_2\right) \\
&= iI_{(4)} \otimes (i\sigma_3) = \begin{pmatrix} -I_{(4)} & 0_{(4)} \\ 0_{(4)} & +I_{(4)} \end{pmatrix}.
\end{aligned}
$$

Let us now examine the chiral spinors in more detail. Consider the left-handed chiral spinor $\Psi_L \equiv P_L\Psi$, where

$$
P_L \equiv \tfrac{1}{2}(I - \gamma^{\text{FIVE}}) = \begin{pmatrix} \tfrac{1}{2}(I_{(4)} + \gamma^{\text{FIVE}}_{(4)}) & 0_{(4)} \\ 0_{(4)} & \tfrac{1}{2}(I_{(4)} - \gamma^{\text{FIVE}}_{(4)}) \end{pmatrix} \equiv \begin{pmatrix} p_R & 0 \\ 0 & p_L \end{pmatrix}.
$$

The 8-dimensional spinor Ψ is therefore projected to a 4-dimensional spinor, as expected, but the components are arranged slightly nontrivially. Write $\Psi = \begin{pmatrix} \psi \\ \psi' \end{pmatrix}$, where ψ and ψ' transform as $(3+1)$-dimensional Dirac spinors, which in turn reduce into two 2-dimensional chiral spinors:

$$
\psi = \begin{pmatrix} \chi_\alpha \\ \bar{\chi}^{\dot\alpha} \end{pmatrix}, \quad \psi' = \begin{pmatrix} \chi'_\alpha \\ \bar{\chi}'^{\dot\alpha} \end{pmatrix}.
$$

In terms of these, the $SO(6)$ left-handed chiral spinor is

$$\Psi_L = \begin{pmatrix} p_R\psi \\ p_L\psi' \end{pmatrix} = \begin{pmatrix} 0 \\ \bar{\chi}^{\dot{\alpha}} \\ \chi'_\beta \\ 0 \end{pmatrix} .$$

So if we want to extract the parts of the rotation generators $\sigma^{\mu\nu}$ that act only on these components, then we may consider the matrices $\sigma_L^{\mu\nu} \equiv P_L\sigma^{\mu\nu}P_L$.

It is now helpful to write the Dirac matrices for $(3 + 1)$-dimensional spacetime in terms of the Pauli matrices:

$$\gamma_{(4)}^\mu = \begin{pmatrix} 0 & (\sigma^\mu)_{\alpha\dot{\beta}} \\ (\bar{\sigma}^\mu)^{\dot{\alpha}\beta} & 0 \end{pmatrix}$$

where numerically we have $(\sigma^\mu)_{\alpha\dot{\beta}} = (I, i\vec{\sigma})$ and $(\bar{\sigma}^\mu)^{\dot{\alpha}\beta} \equiv \varepsilon^{\alpha\beta}\varepsilon^{\dot{\alpha}\dot{\beta}}(\sigma^\mu)_{\beta\dot{\beta}} = (I, -i\vec{\sigma})$.

For $\mu, \nu = 0, 1, 2, 3$, we have:

$$\sigma^{\mu\nu} = \begin{pmatrix} \sigma_{(4)}^{\mu\nu} & 0_{(4)} \\ 0_{(4)} & \sigma_{(4)}^{\mu\nu} \end{pmatrix}$$

$$= \begin{pmatrix} i\frac{1}{2}(\sigma^\mu\bar{\sigma}^\nu - \sigma^\nu\bar{\sigma}^\mu) & 0_{(2)} & 0_{(2)} & 0_{(2)} \\ 0_{(2)} & i\frac{1}{2}(\bar{\sigma}^\mu\sigma^\nu - \bar{\sigma}^\nu\sigma^\mu) & 0_{(2)} & 0_{(2)} \\ 0_{(2)} & 0_{(2)} & i\frac{1}{2}(\sigma^\mu\bar{\sigma}^\nu - \sigma^\nu\bar{\sigma}^\mu) & 0_{(2)} \\ 0_{(2)} & 0_{(2)} & 0_{(2)} & i\frac{1}{2}(\bar{\sigma}^\mu\sigma^\nu - \bar{\sigma}^\nu\sigma^\mu) \end{pmatrix} .$$

Therefore, the middle 4-by-4 block of $\sigma_L^{\mu\nu}$ (for $\mu, \nu = 0, 1, 2, 3$ only) is given by the matrix

$$W^{\mu\nu} = \begin{pmatrix} i\frac{1}{2}(\bar{\sigma}^\mu\sigma^\nu - \bar{\sigma}^\nu\sigma^\mu) & 0_{(2)} \\ 0_{(2)} & i\frac{1}{2}(\sigma^\mu\bar{\sigma}^\nu - \sigma^\nu\bar{\sigma}^\mu) \end{pmatrix} .$$

The middle 4-by-4 block of $\sigma_L^{\mu 4}$ (for $\mu = 0, 1, 2, 3$ only) is given by the matrix

$$X^\mu = \begin{pmatrix} 0_{(2)} & +i\bar{\sigma}^\mu \\ -i\sigma^\mu & 0_{(2)} \end{pmatrix}$$

and that of $\sigma_L^{\mu 5}$ is given by

$$Y^\mu = \begin{pmatrix} 0_{(2)} & \bar{\sigma}^\mu \\ \sigma^\mu & 0_{(2)} \end{pmatrix} .$$

Finally, the middle 4-by-4 block of σ_L^{45} is given by

$$Z = \begin{pmatrix} -I_{(2)} & 0_{(2)} \\ 0_{(2)} & +I_{(2)} \end{pmatrix} .$$

The $\frac{1}{2}(4 \times 3) = 6$ matrices $W^{\mu\nu}$, the 4 matrices X^μ, the 4 matrices Y^μ and the matrix Z constitute $6+4+4+1 = 15$ traceless hermitian 4-by-4 matrices. These therefore generate $SU(4)$,

264

and we have proven that $SO(6)$ and $SU(4)$ are locally isomorphic.

We have shown explicitly that the 4-dimensional representation of $SU(4)$ is equivalent to the 6-dimensional representation of $SO(6)$; in other words, the spinor of $SO(6)$ is the fundamental of $SU(4)$. Now we identify a few higher dimensional representations.

Let $\psi_A \sim 4$ and $\psi^A \sim \bar{4}$ of $SU(4)$. Then $\psi_{[AB]} \sim 4 \otimes_A 4 \sim \frac{4 \times 3}{2} = 6$ of $SU(4)$. Let $\psi^\mu \sim 6$ of $SO(6)$. We have shown that the spinor of $SO(6)$ is the fundamental of $SU(4)$, and now we discover that the antisymmetric tensor of $SU(4)$ is the fundamental of $SO(6)$. Explicitly, we can construct a symbol $\Sigma^\mu_{[AB]}$ out of the familiar Pauli matrices which is invariant under $SO(6) \simeq SU(4)$ transformations on all of its indices simultaneously.[35]

We also have $\psi_{(AB)} \sim 4 \otimes_S 4 \sim \frac{4 \times 5}{2} = 10$ of $SU(4)$. Recall that $SU(4)$ has the 4-index invariant tensor ε^{ABCD}, so that 6 and $\bar{6}$ are equivalent representations (that is, two lower indices are group-equivalent to two upper indices): $\psi^{AB} = \frac{1}{2}\varepsilon^{ABCD}\psi_{CD}$.

We also have $\psi_A{}^B \sim 4 \otimes \bar{4} = 1 \oplus 15$, where the singlet is the trace $\operatorname{tr}\psi = \delta_B{}^A \psi_A{}^B$ and the 15 is comprised of the traceless hermitian matrices displayed previously, or in other words the adjoint representation of $SU(4)$. Also $\psi^{[\mu\nu]} \sim 6 \otimes_A 6 \sim \frac{6 \times 5}{2} = 15$ of $SO(6)$, which is the set of traceless antisymmetric 6-by-6 matrices and therefore the adjoint of $SO(6)$. Indeed, the adjoint of $SO(6)$ is equivalent to the adjoint of $SU(4)$, as we have just proven by showing that the two groups are locally isomorphic.

6. Show that (unfortunately) the number of families we get in $SO(18)$ depends on which subgroup of $SO(8)$ we take to be hypercolor.

Solution:

Let us use the notation S_R to denote the "right-handed" spinor S, by which we mean the chiral spinor with eigenvalue $+1$ under the γ^{FIVE} matrix, and S_L to denote the chiral spinor with eigenvalue -1 under the γ^{FIVE} matrix. As shown in the reference for problem VII.7.3 (equations $A20a, b$), the chiral spinors 2_R^{n+m-1} and 2_R^{n+m-1} of $SO(2n+2m)$ decompose under restriction to the subgroup $SO(2n) \otimes SO(2m)$ as

$$2_R^{n+m-1} \to (2_R^{n-1}, 2_R^{m-1}) \oplus (2_L^{n-1}, 2_L^{m-1})$$
$$2_L^{n+m-1} \to (2_R^{n-1}, 2_L^{m-1}) \oplus (2_L^{n-1}, 2_R^{m-1})$$

For $n = 5$, $m = 4$ these become

$$256_R \to (16_R, 8_R) \oplus (16_L, 8_L)$$
$$256_L \to (16_R, 8_L) \oplus (16_L, 8_R)$$

---

[35]See the appendix of http://arxiv.org/pdf/0902.0981.

for the restriction of $SO(18)$ to the subgroup $SO(10) \otimes SO(8)$. Let us choose the 256_R as the spinor in which to place all of the fermions. We want a subgroup G_{HC} of $SO(8)$ such that 8_L decomposes into a bunch of stuff that is confined, meaning non-singlets of G_{HC}, while 8_R decomposes into a bunch of free particles that do not participate in these strong "hypercolor" interactions, meaning we want 8_R to yield singlets of G_{HC}.

Upon restriction to the special unitary subgroup $SU(4)$ of $SO(8)$, we have

$$8_R \to [0] \oplus [2] \oplus [4] = [0] \oplus [2] \oplus [0]$$
$$8_L \to [1] \oplus [3] = [1] \oplus [\bar{1}]$$

where $[k]$ means the antisymmetric tensor with k indices of the 4-representation (the defining representation) of $SU(4)$. Since $SU(4)$ has the invariant tensor ε_{ABCD} (where A, B, C, D are 4-rep indices, and therefore each runs from 1 to 4), the antisymmetrized 4-index tensor $[4]$ is a singlet under $SU(4)$. That is, $\psi^{[ABCD]} \sim \varepsilon_{ABCD} \psi^{ABCD} \sim 1$ of $SU(4)$. Similarly, $\psi^{[ABC]} \sim_{DABC} \psi^{ABC} \sim \psi_D \sim 4$ of $SU(4)$, or in other words $[3] \sim [\bar{1}]$. (We use the $\sim$ notation to mean "transforms as".)

In any case, we see that the restriction $SO(8) \to SU(4)$ exhibits the desired properties: 8_R contains $SU(4)$-singlets, while 8_L does not. At this stage, we see that the $SO(10) \otimes SU(4)$ spinors $(16_R, [0])_1$ and $(16_R, [0])_2$ (where "1" and "2" are merely labels) constitute two copies of the spinor S^+ of equation (14) on p. 425 of the text – that is, we have discovered exactly two families of fermions, which upon the decomposition $SO(10) \to SU(5) \to SU(3)_c \otimes SU(2)_W \otimes U(1)_Y$ become two families of the low-energy matter fields of the Standard Model.

Since in fact we observe three families, the theory as it stands with $G_{HC} = SU(4)$ is wrong, or incomplete. We must break $SU(4)$ down further. Let ψ_A transform as a 4 of $SU(4)$. Suppose we decide to restrict $SU(4)$ to the subgroup $SU(2) \otimes SU(2)$, meaning

$$\begin{pmatrix} 4 \times 4 \text{ special unitary matrix} \end{pmatrix} \to \begin{pmatrix} 2 \times 2 \text{ special unitary matrix} & O \\ O & 2 \times 2 \text{ special unitary matrix} \end{pmatrix}$$

The index $A = 1, 2, 3, 4$ thus breaks up into an index $a = 1, 2$ to denote transformation under the upper 2×2 block, and an index $\dot{a} = 1, 2$ to denote transformation under the lower 2×2 block. Thus under the restriction $SU(4) \to SU(2) \otimes SU(2)$, we have $\psi_A \to \psi_a \oplus \psi_{\dot{a}}$. In other words, $[1] \to [1, 0] \oplus [0, 1]$.

Now consider the antisymmetric two-index tensor $\psi_{[AB]} \sim 4 \otimes_A 4$ of $SU(4)$. If we restrict to $SU(2) \otimes SU(2)$ in exactly the same way, we see $\psi_{[AB]} \to \psi_{[ab]} \oplus \psi_{a\dot{a}} \oplus \psi_{[\dot{a}\dot{b}]}$, or in other words $[2] \to [2, 0] \oplus [1, 1] \oplus [0, 2]$. However, recall that $SU(2)$ has the invariant tensor ε^{ab}, so actually $\psi_{[ab]} \sim \varepsilon^{ab} \psi_{ab} \sim 1$ is invariant under $SU(2)$. Thus $[2] \to [0, 0] \oplus [1, 1] \oplus [0, 0]$, and we have discovered two more singlets! Therefore, if we take $G_{HC} = SU(2) \otimes SU(2)$, the spinors

$$(16_R, [0])_1, \ (16_R, [0])_2, \ (16_R, [0])_3, \ (16_R, [0])_4$$

266

where the old guys "1" and "2" appeared at $SO(8) \to SO(4)$, while the new guys "3" and "4" appeared at $SO(4) \to SU(2) \otimes SU(2)$, comprise *four* families of $SO(10)$ spinors. With $G_{\mathrm{HC}} = SU(2) \otimes SU(2) \subset SO(8)$, we predict the three observed fermion families plus an additional as of yet unobserved fourth generation.

At this point we see how the game is played and simply quote from section III of the reference: "...the number of $V - A$ families is 'predicted' to be two, three, four, or five, respectively, according to whether $SO(8)$ is broken down to $SO(6)$, $SO(5)$, $SO(4)$, or $SO(3)$. [These orthogonal subgroups are not embedded in $SO(8)$ in the obvious way, however.]" For further details, consult the literature.

7. If you want to grow up to be a string theorist, you need to be familiar with the Dirac equation in various dimensions but especially in 10. As a warm up, study the Dirac equation in 2-dimensional spacetime. Then proceed to study the Dirac equation in 10-dimensional spacetime.

Solution:

The salient feature will be whether the solutions to the Dirac equation can be classified as Weyl, Majorana, or both.[36] In general, let Ψ denote a Dirac spinor in any dimension. Then a Weyl spinor is one that satisfies $\Psi = \Psi_{L \text{ or } R}$, where L, R denote chiral projections, and a Majorana spinor is one that satisfies $\Psi = \Psi^c$, where ^{c} denotes the charge conjugation operation in whichever dimension is relevant.

First consider $d = 1 + 1$. The gamma matrices can be taken as

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} , \quad \gamma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

which satisfy the Clifford algebra $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$ with $\eta = \mathrm{diag}(-, +)$. The Dirac operator $i\partial\!\!\!/ \equiv i\gamma^\mu \partial_\mu$ is diagonal:

$$i\partial\!\!\!/ = - \begin{pmatrix} \partial_0 - \partial_1 & 0 \\ 0 & \partial_0 + \partial_1 \end{pmatrix}$$

and therefore, writing $\Psi = \begin{pmatrix} \psi \\ \tilde{\psi} \end{pmatrix}$, we see that the two spinors $\psi, \tilde{\psi}$ do not mix under the Dirac equation and therefore can be taken as independent degrees of freedom.

If we define the chiral projector $\gamma^{\mathrm{FIVE}} \equiv \gamma^0 \gamma^1 = \mathrm{diag}(+1, -1)$, then we can define the chiral spinors $\Psi_L \equiv \frac{1}{2}(I - \gamma^{\mathrm{FIVE}})\Psi = \begin{pmatrix} 0 \\ \tilde{\psi} \end{pmatrix}$ and $\Psi_R \equiv \frac{1}{2}(I + \gamma^{\mathrm{FIVE}})\Psi = \begin{pmatrix} \psi \\ 0 \end{pmatrix}$, which satisfy

---

[36]We follow http://www.kitp.ucsb.edu/ joep/JLBS.pdf as well Appendix B of *String Theory, Volume II* by J. Polchinski.

$\gamma^{\mathrm{FIVE}}\Psi_L = (-1)\Psi_L$ and $\gamma^{\mathrm{FIVE}}\Psi_R = (+1)\Psi_R$.

The Lorentz generators in the spinor representation are $\sigma^{\mu\nu} \equiv -i\frac{1}{4}[\gamma^\mu, \gamma^\nu]$, which in our case reduces to $\sigma^{01} = -i\frac{1}{2}\gamma^0\gamma^1 \propto \gamma^{\mathrm{FIVE}}$, so that $[\gamma^{\mathrm{FIVE}}, \sigma^{01}] = 0$. Therefore the spinors $\Psi_{L,R}$ transform irreducibly under the Lorentz group.

Therefore, given a Dirac spinor Ψ, we may impose a Weyl condition $\Psi = \Psi_L$ or $\Psi = \Psi_R$ in two spacetime dimensions.

Furthermore, since it is possible to choose a basis in which the Dirac operator is purely real (as we have done), it is consistent to impose that its solutions are real: $\psi^* = \psi$, $\tilde{\psi}^* = \tilde{\psi}$, or in other words $\Psi^* = \Psi$.

Thus, if we define the charge-conjugated Dirac field $\Psi^c \equiv \Psi^*$ as well as the left-chiral Dirac field $\Psi_L \equiv \frac{1}{2}(I - \gamma^{\mathrm{FIVE}})\Psi$, then we find it is consistent to impose the condition $\Psi = \Psi^c = \Psi_L$. A spinor satisfying this condition is called a Majorana-Weyl spinor. (We can also use Ψ_R instead of Ψ_L.)

Before jumping to $d = 9 + 1$, let us compare this briefly with what we know from $d = 3 + 1$. We learned in Chapter II (and as explained in Appendix E), a Dirac spinor Ψ in $(3 + 1)$-dimensional spacetime transforms reducibly as $4_{\mathrm{Dirac}} = 2_L \oplus 2_R$:

$$\Psi = \begin{pmatrix} \psi_\alpha \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix}$$

where $\psi \sim (2, 1)$ and $\bar{\chi} \sim (1, 2)$ of $SU(2)_L \otimes SU(2)_R \simeq SO(3, 1)$. It is therefore consistent to impose the Weyl condition $\Psi = \Psi_L = \begin{pmatrix} \psi_\alpha \\ 0 \end{pmatrix}$ or $\Psi = \Psi_R = \begin{pmatrix} 0 \\ \bar{\chi}^{\dot{\alpha}} \end{pmatrix}$ using the gamma matrix conventions given in Chapter II.

However, we also learned that in $(3 + 1)$-dimensional spacetime, hermitian conjugation interchanges the two chiral representations. In other words, $\psi_\alpha \sim (2, 1) \implies (\psi_\alpha)^\dagger = (\psi^\dagger)_{\dot{\alpha}} \sim (1, 2)$, and $\bar{\chi}^{\dot{\alpha}} \sim (1, 2) \implies (\bar{\chi}^{\dot{\alpha}})^\dagger = (\bar{\chi}^\dagger)^\alpha \sim (2, 1)$. (In other words, $2_R = \bar{2}_L$.) Therefore it is not possible to impose a condition such as "$\psi^\dagger = \psi$" or "$\bar{\chi}^\dagger = \bar{\chi}$" in $d = 3 + 1$.

However, it is possible to impose the condition $\bar{\chi}^\dagger = \psi$. In other words, if we define the conjugate spinor

$$\Psi^c \equiv \begin{pmatrix} (\bar{\chi}^\dagger)_\alpha \\ (\psi^\dagger)^{\dot{\alpha}} \end{pmatrix}$$

then it is possible to impose the Majorana condition $\Psi = \Psi^c$.

The key point is that it is not possible to impose both the Weyl and Majorana conditions simultaneously. For example, if $\Psi = \Psi_L = \begin{pmatrix} \psi \\ 0 \end{pmatrix}$, then $\Psi^c = \begin{pmatrix} 0 \\ \psi^\dagger \end{pmatrix} \neq \Psi$. Conversely, if

$\Psi = \Psi^c = \begin{pmatrix} \psi \\ \psi^\dagger \end{pmatrix}$, then $\Psi_L = \begin{pmatrix} \psi \\ 0 \end{pmatrix} \neq \Psi$. In $d = 3+1$, a Weyl spinor cannot be Majorana, and a Majorana spinor cannot be Weyl.

Now consider $d = 9+1$. As explained in the text, the gamma matrices can be constructed iteratively starting from $d = 1+1$:

$$\gamma^0_{(d=2)} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \gamma^1_{(d=2)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\gamma^{\mu=0,\dots,7}_{(d=10)} = \gamma^\mu_{(d=8)} \otimes \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \gamma^8_{(d=10)} = I_{16} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^9_{(d=10)} = I_{16} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

where I_{16} is the $16 = 2^4$-dimensional identity matrix. (The gamma matrices have size $2^{d/2}$, which for $d = 10$ is $2^5 = 32$. For instance, the matrix $I_{16} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is indeed 32-by-32.)

Define the chiral projection matrix $\gamma^{\text{FIVE}} \equiv \gamma^0 \gamma^1 \gamma^2 \dots \gamma^9$. The chiral spinor $\Psi_L \equiv \frac{1}{2}(1 - \gamma^{\text{FIVE}})\Psi$ satisfies $\gamma^{\text{FIVE}}\Psi_L = (-1)\Psi_L$, and the chiral spinor $\Psi_R \equiv \frac{1}{2}(1 - \gamma^{\text{FIVE}})\Psi$ satisfies $\gamma^{\text{FIVE}}\Psi_R = (+1)\Psi_R$.

The question now is whether these irreducible chiral representations are self conjugate or conjugate to each other. Under Lorentz transformations, the Dirac spinor Ψ transformas as

$$\Psi \to e^{i\omega_{\mu\nu}\sigma^{\mu\nu}}\Psi = (I + i\omega_{\mu\nu}\sigma^{\mu\nu} + \dots)\Psi$$

where $\sigma^{\mu\nu} \equiv -i\frac{1}{4}[\gamma^\mu, \gamma^\nu]$ generates Lorentz transformations in the $2^5 = 32$-dimensional (reducible) spinor representation of $SO(9,1)$. If we can find a matrix B such that $B\sigma^{\mu\nu}B^{-1} = -(\sigma^{\mu\nu})^*$, then the spinor $B^{-1}\Psi^*$ transforms under infinitesimal Lorentz transformations as:

$$\begin{aligned} B^{-1}\Psi^* \to \ & B^{-1}(I + i\omega_{\mu\nu}\sigma^{\mu\nu})^*\Psi^* \\ &= B^{-1}(I - i\omega_{\mu\nu}(\sigma^{\mu\nu})^*)\Psi^* \\ &= (I + i\omega_{\mu\nu}\sigma^{\mu\nu})B^{-1}\Psi^* \end{aligned}$$

which is identical to the transformation property of Ψ. Thus the (reducible) Dirac representation Ψ is self-conjugate. The question is whether its irreducible chiral components are also self-conjugate.

By explicitly anticommuting the gamma matrices, we see that the matrix $B = \gamma^3\gamma^5\gamma^7\gamma^9$ has the desired property: $B\sigma^{\mu\nu}B^{-1} = -(\sigma^{\mu\nu})^*$. Now that we have an explicit form for B, we can compute its effect on the chiral projection matrix γ^{FIVE}: $B\gamma^{\text{FIVE}}B^{-1} = +\gamma^{\text{FIVE}}$. Therefore, if $\Psi = \Psi_L$ so that $\gamma^{\text{FIVE}}\Psi = (-1)\Psi$, then $\Psi_L^B \equiv B^{-1}\Psi_L^*$ is also left-handed: $\gamma^{\text{FIVE}}\Psi_L^B = (-1)\Psi_L^B$. Thus the left-handed chiral spinor is self-conjugate. Similarly, the right-handed chiral spinor is also self-conjugate.

With an extra factor of γ^0, we define the charge conjugation matrix $C \equiv B\gamma^0$ so that

$\Psi^c \equiv C\bar{\Psi}^T = B\Psi^*$ denotes the charge conjugate spinor. We have therefore learned that, just like $d = 1+1$ but unlike $d = 3+1$, it is consistent to impose the condition $\Psi = \Psi_L = \Psi^c$ in $d = 9+1$ dimensions. Again, a spinor satisfying this condition is called a Majorana-Weyl spinor.

VIII   Gravity and Beyond

VIII.1   Gravity as a Field Theory and the Kaluza-Klein Picture

1. Work out $T^{\mu\nu}$ for a scalar field. Draw the Feynman diagram for the contribution of one-graviton exchange to the scattering of two scalar mesons. Calculate the amplitude and extract the interaction energy between two mesons sitting at rest, thus deriving Newton's law of gravity.

Solution:

For this and the next problem, we will need to vary the determinant of the metric.

$$\det(g + \delta g) = \det g \det(1 + g^{-1}\delta g) = \det g \, e^{\text{tr} \, \ln(1 + g^{-1}\delta g)}$$
$$= \det g \, e^{\text{tr}(g^{-1}\delta g)} = \det g \left[1 + \text{tr}(g^{-1}\delta g) \right]$$

So defining as usual $\delta \det g = \det(g + \delta g) - \det g$, we get $\delta \det g = \det g \, \text{tr}(g^{-1}\delta g)$. It will be convenient to vary with respect to g^{-1} rather than g, so to relate the two consider the definition $gg^{-1} = 1$. Varying this yields $\delta gg^{-1} + g\delta g^{-1} = 0 \implies \delta g = -g \, \delta g^{-1}g$. Therefore we have

$$\delta \det g = - \det g \, \text{tr}(\delta g^{-1}g)$$

We actually need $\delta\sqrt{-\det g}$, which is:

$$\delta\sqrt{-\det g} = \frac{1}{2\sqrt{-\det g}} \delta \det g = -\frac{1}{2}\sqrt{-\det g} \, \text{tr}(\delta g^{-1} g)$$

Just to clarify, $\text{tr}(\delta g^{-1} g) = \delta g^{\mu\nu}g_{\mu\nu}$. Now we are ready for the problem. The action for a real scalar field ϕ is

$$S = \int d^4x \sqrt{-\det g} \, \mathcal{L} \, , \quad \mathcal{L} = \frac{1}{2}g^{\mu\nu}\partial_\mu \phi \, \partial_\nu \phi - V(\phi) \, .$$

Treating this as a function of g^{-1}, the variation $\delta S \equiv S[g^{-1} + \delta g^{-1}] - S[g^{-1}]$ is

$$\delta S = \int d^4x \left[\delta\sqrt{-\det g} \, \mathcal{L} + \sqrt{-\det g} \, \delta\mathcal{L} \right]$$
$$= \int d^4x \sqrt{-\det g} \left[-\frac{1}{2}g_{\mu\nu}\mathcal{L} + \frac{1}{2}\partial_\mu \phi \, \partial_\nu \phi \right] \delta g^{\mu\nu} \, .$$

So the stress tensor is $T_{\mu\nu} \equiv \frac{2}{\sqrt{-\det g}}\delta S/\delta g^{\mu\nu} = \partial_\mu \phi \, \partial_\nu \phi - g_{\mu\nu}\mathcal{L}$ with the Lagrangian $\mathcal{L} = \frac{1}{2}g^{\mu\nu}\partial_\mu \phi \, \partial_\nu \phi - V(\phi)$.

For a free field theory, $V(\phi) = \frac{1}{2}m^2\phi^2$, so the stress tensor is

$$T_{\mu\nu} = \partial_\mu \phi \, \partial_\nu \phi - \frac{1}{2}g_{\mu\nu}\left(g^{\rho\sigma}\partial_\rho \phi \, \partial_\sigma \phi - m^2\phi^2\right) \, .$$
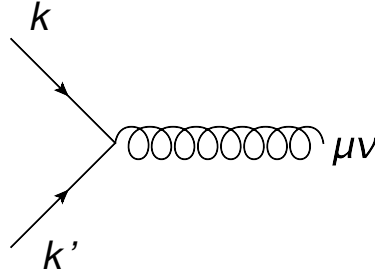
Now we need the weak field action for gravity coupling to matter. This is given in equation (10) on p. 437, which we repeat below:

$$S = \int d^4x \, \frac{1}{2} \left[\frac{1}{32\pi G} \left(\partial_\lambda h^{\mu\nu} \partial^\lambda h_{\mu\nu} - \tfrac{1}{2} \partial_\lambda h^\mu_\mu \partial^\lambda h^\nu_\nu \right) - h^{\mu\nu} T_{\mu\nu} \right]$$

Define $M_P \equiv 1/\sqrt{16\pi G}$ and rescale the graviton as $h^{\mu\nu} \to \frac{1}{M_P} \sqrt{2} \, h^{\mu\nu}$ to obtain

$$S = \int d^4x \, \left\{ \tfrac{1}{2} \left(\partial_\lambda h^{\mu\nu} \partial^\lambda h_{\mu\nu} - \tfrac{1}{2} \partial_\lambda h^\mu_\mu \partial^\lambda h^\nu_\nu \right) - \frac{1}{\sqrt{2} \, M_P} h^{\mu\nu} T_{\mu\nu} \right\}$$

We have the interaction term $S_{\text{int}} = -\frac{1}{\sqrt{2} M_P} \int d^4x \, T^{\mu\nu} h_{\mu\nu}$, which leads to the momentum space vertex represented pictorially as

which is equal to

$$\frac{-i}{\sqrt{2} \, M_P} \left[k^\mu k'^\nu + k^\nu k'^\mu - \eta^{\mu\nu} (k \cdot k' + m^2) \right]$$

where both momenta flow into the vertex.

We also have the graviton propagator

which, in the harmonic gauge, is equal to

$$i\Delta_{\mu\nu,\lambda\sigma}(k) \equiv \frac{i}{2k^2} \left(\eta_{\mu\lambda}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\lambda} - \eta_{\mu\nu}\eta_{\lambda\sigma} \right) .$$

Consider the case of two different scalar fields ϕ_1 and ϕ_2, each with its own stress tensor. We want to compute the gravitational potential between a ϕ_1 particle and a ϕ_2 particle both sitting at rest, or in other words in the extreme nonrelativistic limit. In this limit, only the t-channel diagram for $\phi_1\phi_2 \to \phi_1\phi_2$ 2-to-2 scattering contributes. Also, the 00-component of the stress tensor dominates, so the vertex becomes

$$\frac{-i}{\sqrt{2} \, M_P} \left(k^0 k'^0 + \vec{k} \cdot \vec{k}' - m^2 \right) \delta^\mu_0 \delta^\nu_0$$

The amplitude is $i\mathcal{M} =$

272
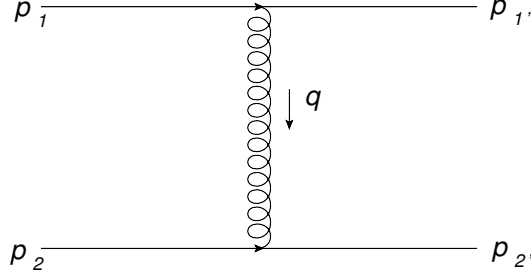
$$= \frac{-i}{\sqrt{2}\,M_P}\left[-p_1^0 p_{1'}^0 - \vec{p}_1\cdot\vec{p}_{1'} - m_1^2\right]\left[i\Delta_{00,00}(q)\right]\frac{-i}{\sqrt{2}\,M_P}\left[-p_2^0 p_{2'}^0 - \vec{p}_2\cdot\vec{p}_{2'} - m_2^2\right]$$

$$= \frac{-i}{2^2 M_P^2 q^2}(p_1^0 p_{1'}^0 + \vec{p}_1\cdot\vec{p}_{1'} + m_1^2)(p_2^0 p_{2'}^0 + \vec{p}_2\cdot\vec{p}_{2'} + m_2^2)$$

where $q = p_1 - p_{1'} = p_{2'} - p_2$ is the momentum transfer. The minus signs on the momenta in the second line come from the fact that in the diagram, one momentum flows into the vertex while the other flows out. In the definition of the vertex given above, both momenta flow into the vertex.

If the two incoming particles are at rest, then $p_1^\mu = (m_1, \vec{0})$ and $p_2^\mu = (m_2, \vec{0})$, which implies $(p_1^0 p_{1'}^0 + \vec{p}_1\cdot\vec{p}_{1'} + m_1^2)(p_2^0 p_{2'}^0 + \vec{p}_2\cdot\vec{p}_{2'} + m_2^2) = (2m_1^2)(2m_2^2) = 4m_1^2 m_2^2$. In the denominator we have

$$q^2 = (p_1 - p_{1'})^2 = p_1^2 + p_{1'}^2 - 2p_1\cdot p_{1'}$$

$$= m_1^2 + m_1^2 - 2m_1\sqrt{|\vec{p}_{1'}|^2 + m_1^2}$$

$$= 2m_1^2 - 2m_1^2\left[1 + \frac{|\vec{p}_{1'}|^2}{2m_1^2} + O\left(\frac{|\vec{p}_{1'}|^4}{m_1^4}\right)\right]$$

$$= -|\vec{p}_{1'}|^2 + (\text{higher order})$$

The amplitude is

$$\mathcal{M} = \frac{m_1^2 m_2^2}{M_P^2 |\vec{p}_{1'}|^2} .$$

Scattering from a potential in nonrelativistic quantum mechanics gives[37]

$$V(\vec{x}) = -\frac{1}{4m_1 m_2}\int \frac{d^3q}{(2\pi)^3}\, e^{i\vec{q}\cdot\vec{x}}\mathcal{M}(\vec{q})$$

where $\vec{q} = \vec{p}_1 - \vec{p}_{1'} = -\vec{p}_{1'}$ is the transfer of 3-momentum. Since

$$\int \frac{d^3q}{(2\pi)^3}\, e^{i\vec{q}\cdot\vec{x}}\frac{1}{|\vec{q}|^2} = \frac{1}{4\pi|\vec{x}|}$$

---

[37]See Section 4.5.7 of http://arxiv.org/abs/0812.1594 and the textbook M. Maggiore, A Modern Introduction to Quantum Field Theory (Oxford University Press, Oxford, UK, 2005) pp. 167-170 .

we have

$$V(\vec{x}) = -\frac{1}{4m_1m_2}\left(\frac{m_1^2m_2^2}{M_P^2}\right)\left(\frac{1}{4\pi|\vec{x}|}\right) = -\frac{m_1m_2}{16\pi M_P^2|\vec{x}|} \ .$$

The Planck mass is defined in terms of Newton's constant as $M_P^2 = 1/(16\pi G)$, so we get

$$V(\vec{x}) = -\frac{Gm_1m_2}{|\vec{x}|}$$

which is Newton's inverse square law.

Note that various conventions exist for defining the Planck mass M_P in terms of Newton's constant G, and so various factors of 2 can appear between references. For a discussion of graviton scattering in various contexts, see K. Hinterbichler, "Theoretical aspects of massive gravity," arXiv:1105.3735v1 [hep-th].

2. Work out $T^{\mu\nu}$ for the Yang-Mills field.

Solution:

The Lagrangian for the Yang-Mills field is

$$\mathcal{L} = -\frac{1}{2}\mathrm{tr}(F_{\mu\nu}F^{\mu\nu}) = -\frac{1}{2}g^{\mu\rho}g^{\nu\sigma}\mathrm{tr}(F_{\mu\nu}F_{\rho\sigma}) = -\frac{1}{4}g^{\mu\rho}g^{\nu\sigma}F^a_{\mu\nu}F^a_{\rho\sigma}$$

Varying the action $S = \int d^4x\sqrt{-\det g}\,\mathcal{L}$ gives with respect to the inverse metric gives

$$\delta S = -\frac{1}{4}\int d^4x\,\sqrt{-\det g}\left(-\frac{1}{2}g_{\alpha\beta}F^a_{\mu\nu}F^{a\mu\nu} + g^{\nu\sigma}F^a_{\alpha\nu}F^a_{\beta\sigma} + g^{\mu\rho}F^a_{\mu\alpha}F^a_{\rho\beta}\right)\delta g^{\alpha\beta}$$

From the antisymmetry $F^a_{\mu\nu} = -F^a_{\nu\mu}$ the last two terms are the same, so we get

$$T^{\mathrm{Yang\text{-}Mills}}_{\alpha\beta} = g^{\mu\nu}F^a_{\alpha\mu}F^a_{\beta\nu} - \frac{1}{4}g_{\alpha\beta}F^a_{\mu\nu}F^{a\mu\nu}$$

Note that since $\gamma^{\alpha\beta}g^{\mu\nu}F^a_{\alpha\mu}F^a_{\beta\nu} = F^a_{\mu\nu}F^{a\mu\nu}$ and $g^{\alpha\beta}g_{\alpha\beta} = 4$, this stress tensor is in fact traceless.

3. Show that if $h_{\mu\nu}$ does not satisfy the harmonic gauge, we can always make a gauge transformation with ε_ν determined by $\partial^2 \varepsilon_\nu = \partial_\mu h_\nu^\mu - \frac{1}{2}\partial_\nu h_\lambda^\lambda$ so that it does. All of this should be conceptually familiar from your study of electromagnetism.

Solution:

Using the gauge transformation $h_{\mu\nu} = h'_{\mu\nu} + \partial_\mu\varepsilon_\nu + \partial_\nu\varepsilon_\mu$, we have:

$$\partial_\mu h_\nu^\mu - \frac{1}{2}\partial_\nu h_\lambda^\lambda = \partial_\mu(h_\nu'^\mu + \partial^\mu\varepsilon_\nu + \partial_\nu\varepsilon^\mu) - \frac{1}{2}\partial_\nu(h_\lambda'^\lambda + 2\partial_\lambda\varepsilon^\lambda)$$
$$= \partial_\mu h_\nu'^\mu - \frac{1}{2}\partial_\nu h_\lambda'^\lambda + \partial^2\varepsilon_\nu + \partial_\mu\partial_\nu\varepsilon^\mu - \partial_\nu\partial_\lambda\varepsilon^\lambda$$
$$= \partial_\mu h_\nu'^\mu - \frac{1}{2}\partial_\nu h_\lambda'^\lambda + \partial^2\varepsilon_\nu$$

We can choose $\partial^2\varepsilon_\nu = \partial_\mu h_\nu^\mu - \frac{1}{2}\partial_\nu h_\lambda^\lambda$ to make h' satisfy the harmonic gauge.

4. Count the number of degrees of polarization of a graviton. [Hint: Consider a plane wave $h_{\mu\nu}(x) = h_{\mu\nu}(k)e^{ikx}$ just because it is a bit easier to work in momentum space. A symmetric tensor has 10 components and the harmonic gauge $k_\mu h_\nu^\mu = \frac{1}{2}k_\nu h_\lambda^\lambda$ imposes 4 conditions. Oops, we are left with 6 degrees of freedom. What is going on?] [Hint: You can make a further gauge transformation and still stay in the harmonic gauge. The graviton should have only 2 degrees of polarization.]

Solution:

Suppose we have already imposed the condition $k^\mu h_{\mu\nu} = \frac{1}{2}k_\nu h^\lambda_\lambda$ as indicated, so that we have 6 degrees of freedom. Consider the gauge transformation $h_{\mu\nu} = h'_{\mu\nu} - k_\mu\varepsilon_\nu - k_\nu\varepsilon_\mu$. Putting this into the above condition gives

$$k^\mu(h'_{\mu\nu} - k_\mu\varepsilon_\nu - k_\nu\varepsilon_\mu) = \frac{1}{2}k_\nu(h'^\lambda_\lambda - 2k^\lambda\varepsilon_\lambda)$$
$$k^\mu h'_{\mu\nu} - k^2\varepsilon_\nu - k_\nu k^\mu\varepsilon_\mu = \frac{1}{2}k_\nu h'^\lambda_\lambda - k_\nu k^\lambda\varepsilon_\lambda$$
$$k^\mu h'_{\mu\nu} = \frac{1}{2}k_\nu h'^\lambda_\lambda$$

since $k^2 = 0$. The function $h'_{\mu\nu}$ also satisfies the harmonic gauge for arbitrary ε_μ, so we have not completely used up our gauge freedom. Indeed, we still have 4 functions to choose, namely $\{\varepsilon_\mu\}_{\mu=0}^3$, which cuts the number of physical degrees of freedom from 6 to 2. Thus the graviton has 2 physical degrees of freedom.

5. The Kaluza-Klein result that we argued by symmetry considerations can of course be derived explicitly. Let me sketch the calculation for you. Consider the metric

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu - a^2[d\theta + A_\mu(x)dx^\mu]^2$$

where θ denotes an angular variable $0 \le \theta < 2\pi$. With $A_\mu = 0$, this is just the metric of a curved spacetime, which has a circle of radius a attached at every point. The transformation $\theta \to \theta + \Lambda(x)$ leaves ds invariant provided that we also transform $A_\mu(x) \to A_\mu(x) - \partial_\mu\Lambda(x)$. Calculate the 5-dimensional scalar curvature R_5 and show that $R_5 = R_4 - \frac{1}{4}a^2 F_{\mu\nu}F^{\mu\nu}$. Except for the precise coefficient $\frac{1}{4}$ this result follows entirely from symmetry considerations and from the fact that R_5 involves two derivatives on the 5-dimensional metric, as explained in the text. After some suitable rescaling this is the usual action for gravity plus electromagnetism. Note that the 5-dimensional metric has the explicit form

$$g_{AB}^5 = \begin{pmatrix} g_{\mu\nu} - a^2 A_\mu A_\nu & -a^2 A_\mu \\ -a^2 A_\nu & -a^2 \end{pmatrix}.$$

Solution:

We will compute the result using differential forms. First, find the veilbein 1-forms $e^m = e^m_M dx^M = e^m_\mu dx^\mu + e^m_\theta d\theta$, then find the connection 1-forms $\omega^m{}_n$ from $de^m = -\omega^m{}_n e^n$, then compute the curvature 2-form $R^m{}_n = d\omega^m{}_n + \omega^m{}_p \omega^p{}_n$. From the curvature 2-form, extract the components via $R^m{}_n = \frac{1}{2}R^m{}_{nrs}e^r e^s$. Finally, compute the curvature scalar $R \equiv \eta^{ns}R^m{}_{nms}$.

The conventions we use for the indices are:

$$\text{Flat: } m, n, p, ... = 0, 1, 2, 3, 4 \; ; \quad a, b, c, ... = 0, 1, 2, 3$$
$$\text{Curved: } M, N, P, ... = 0, 1, 2, 3, 4 \; ; \quad \mu, \nu, \rho, ... = 0, 1, 2, 3$$

where $x^4 \equiv a\theta$ is the fifth coordinate.

The 5D metric is

$$ds^2 = G_{MN}dx^M dx^N = (g_{\mu\nu} - a^2 A_\mu A_\nu)dx^\mu dx^\nu - a^2 d\theta^2 - aA_\mu(dx^\mu d\theta + d\theta dx^\mu)$$

Defining the veilbeins by

$$G_{MN} = e^m_M \eta_{mn} e^n_N = e^\mu_M \eta_{\mu\nu} e^\nu_N - e^\theta_M e^\theta_N$$

we have:

$$G_{\theta\theta} = -a^2 \implies e^\theta_\theta = a, \; e^a_\theta = 0$$
$$G_{\mu\theta} = -aA_\mu \implies e^\theta_\mu = aA_\mu$$
$$G_{\mu\nu} = g_{\mu\nu} - a^2 A_\mu A_\nu \implies e^a_\mu = e^a_{(4)\mu}$$

Now define the 1-forms $e^m \equiv e^m_M dx^M$. This implies

$$e^a = e^a_{(4)} \; , \quad e^\theta = a(A + d\theta)$$

where $e^a_{(4)}$ are the 4D veilbeins defined by $g_{\mu\nu} = e^a_{(4)}\eta_{ab}e^b_{(4)}$, and $A \equiv A_\mu dx^\mu$. Recall that the point of using forms notation is to be able to expand in any basis we choose, not necessarily the coordinate basis dx^μ. Thus we define the component fields A_a via $A \equiv A_a e^a_{(4)}$, and for later convenience the partial derivatives ∂_a via $dx^\mu \partial_\mu \equiv e^a_{(4)} \partial_a$.

Now we need to find the connection 1-forms ω^m_{n}, which are defined by the equation $de^m + \omega^m_{n}e^n = 0$. Since $e^\theta = aA$, we have[38] $de^\theta = d(d\theta + aA) = 0 + aF = a\frac{1}{2}F_{ab}e^a_{(4)}e^b_{(4)}$. Since $\omega^\theta_{\theta} = 0$, the definition $de^\theta = -\omega^\theta_{m}e^m = -\omega^\theta_{a}e^a_{(4)}$ implies

$$\omega^\theta_{a} = +\frac{1}{2}a\, F_{ab}e^b_{(4)}$$

where we have been careful to keep track of the order of the 1-forms $e^a e^b$ when extracting the components $F_{ab} = -F_{ba}$, hence the $+$ sign in the above.

Next we need $de^a = de^a_{(4)} \equiv -\omega^a_{(4)b}e^b_{(4)}$. We have $de^a \equiv -\omega^a_{m}e^m = -\omega^a_{b}e^b - \omega^a_{\theta}e^\theta$, and[39] $\omega^a_{\theta} = \eta_{\theta\theta}\omega^{a\theta} = -\eta_{\theta\theta}\omega^{\theta a} = -\eta_{\theta\theta}\eta^{ab}\omega^\theta_{b} = -\eta_{\theta\theta}\eta^{ab}(\frac{1}{2}a\,F_{bc}e^c_{(4)}) = -\eta_{\theta\theta}\frac{1}{2}a\,F^a_{c}e^c_{(4)}$. Therefore we have $\omega^a_{(4)c}e^c_{(4)} = \omega^a_{c}e^c_{(4)} - \eta_{\theta\theta}\frac{1}{2}a\,F^a_{c}e^c_{(4)}e^\theta = \omega^a_{c}e^c_{(4)} + \eta_{\theta\theta}\frac{1}{2}a\,F^a_{c}e^\theta e^c_{(4)}$, where we have used the Grassmann property $e^c e^\theta = -e^\theta e^c$. Moving the last term over to the other side of the equation, we can read off the coefficient of $e^c_{(4)}$:

$$\omega^a_{c} = \omega^a_{(4)c} - \frac{1}{2}a\,\eta_{\theta\theta}F^a_{c}e^\theta \ .$$

Now we are ready to define the curvature 2-form, $R^m_{n} = d\omega^m_{n} + (\omega^2)^m_{n}$. First compute $R^a_{b} = d\omega^a_{b} + (\omega^2)^a_{b}$. We have $d\omega^a_{b} = d\omega^a_{(4)b} - \frac{1}{2}a\,\eta_{\theta\theta}d(F^a_{b}e^\theta)$, and $d(F_{ab}e^\theta) = \partial_c F_{ab}e^c_{(4)}e^\theta + F_{ab}(aF) = \partial_c F_{ab}e^c_{(4)}e^\theta + \frac{1}{2}aF_{ab}F_{cd}e^c_{(4)}e^d_{(4)}$. To summarize:

$$d\omega^a_{b} = d\omega^a_{(4)b} - \frac{1}{2}a\,\eta_{\theta\theta}\partial_c F^a_{b}e^c_{(4)}e^\theta - \frac{1}{4}a^2\,\eta_{\theta\theta}F^a_{b}F_{cd}e^c_{(4)}e^d_{(4)} \ .$$

Next we need $(\omega^2)^a_{b} = \omega^a_{m}\omega^m_{b} = \omega^a_{c}\omega^c_{b} + \omega^a_{\theta}\omega^\theta_{b}$. Multiplying the terms out gives

$$(\omega^2)^a_{b} = (\omega^2_{(4)})^a_{b} + \frac{1}{2}a\,\eta_{\theta\theta}(F^a_{c}\omega^c_{(4)b} - \omega^a_{(4)c}F^c_{b})e^\theta - \frac{1}{4}a^2\,F^a_{c}F_{bd}e^c_{(4)}e^d_{(4)} \ .$$

Therefore, the R^a_{b} part of the curvature 2-form is

$$R^a_{b} = R^a_{(4)b} - \frac{1}{4}a^2\,(F^a_{c}F_{bd} + \eta_{\theta\theta}F^a_{b}F_{cd})e^c_{(4)}e^d_{(4)} + \frac{1}{2}a\,\eta_{\theta\theta}(F^a_{c}\omega^c_{(4)b} - \omega^a_{(4)c}F^c_{b} - \partial_c F^a_{b}e^c_{(4)})e^\theta \ .$$

---

[38] To make sure the meaning of the last equality here is clear: In the coordinate basis, we have $F = \frac{1}{2}F_{\mu\nu}dx^\mu dx^\nu$, so that F is a 2-form and the components $F_{\mu\nu}$ are just numbers. We use the definition of the veilbein 1-forms $e^a \equiv e^a_\mu dx^\mu$ to define the c-number components F_{ab} via $F_{\mu\nu} \equiv e^a_\mu e^b_\nu F_{ab}$. So while ω^a_{b} is a collection of 1-forms, $F^a_{b} \equiv \eta^{ac}F_{cb}$ is not a collection of forms but instead is just a collection of numbers.

[39] We will keep the explicit factor of $\eta_{\theta\theta}$ present to show that it cancels out of the final result, and hence that the resulting action is independent of whether we use the metric convention $\eta = (+,-,-,-,-)$ or the one more commonly used in the gravitational literature, $\eta = (-,+,+,+,+)$.

A few comments are in order at this point. First, to compute the curvature scalar we will not need coefficients of the 2-form $e^c_{(4)}e^\theta$, so we will drop them from now on. Second, look at the terms in parentheses multiplying $e^c_{(4)}e^d_{(4)}$. Since $F_{cd} = -F_{dc}$, the second term in parentheses is manifestly antisymmetric in (cd), but the first term is not. Thus to extract the components $R^a{}_{bcd}$ from the 2-form $R^a{}_b = \frac{1}{2}R^a{}_{bcd}e^c e^d$, we must write $F^a{}_c F_{bd}e^c e^d = \frac{1}{2}(F^a{}_c F_{bd} - F^a{}_d F_{bc})e^c e^d$. Therefore, we have

$$R^a{}_{bcd} = (R_{(4)})^a{}_{bcd} - \frac{1}{4}a^2(F^a{}_c F_{bd} - F^a{}_d F_{bc}) - \frac{1}{2}a^2\,\eta_{\theta\theta}F^a{}_b F_{cd}\ .$$

Finally, let us make sure the dimensions are correct. The e^m have dimensions of length (recall $e^m = e^m_\mu dx^\mu$ with e^m_μ dimensionless), so writing $e^\theta = aA$ implies A is dimensionless. Since $A = A_a e^a$, we find A_a has dimensions of inverse length. Since $e^a\partial_a = dx^\mu\partial_\mu$, the derivatives ∂_a have dimensions of inverse length (just like ∂_μ), so $F_{ab} = \partial_a A_b - \partial_b A_a$ has dimensions of $1/\text{length}^2$. Thus $a^2 F^a{}_c F_{bd}$ has dimensions of $1/\text{length}^2$, as it should be to match $R^a{}_{bcd}$. Note that the operator $d = e^a\partial_a$ is dimensionless, so that the 1-forms $\omega^m{}_n$ defined by $de^m = -\omega^m{}_n e^n$ are also dimensionless.

To compute the curvature scalar we still need the 2-form $R^\theta{}_a = d\omega^\theta{}_a + (\omega^2)^\theta{}_a$. Recall that $\omega^\theta{}_a = \frac{1}{2}a\,F_{ab}e^b_{(4)}$, so we have $d\omega^\theta{}_a = \frac{1}{2}a\,d(F_{ab}e^b_{(4)}) = \frac{1}{2}a\,(\partial_c F_{ab}e^c_{(4)}e^b_{(4)} + F_{ab}de^b_{(4)})$. Since $de^b_{(4)} = -\omega^b_{(4)c}e^c_{(4)}$ by definition, we have

$$d\omega^a{}_\theta = \frac{1}{2}a\,\partial_c F_{ab}e^c_{(4)}e^b_{(4)} - \frac{1}{2}a\,F_{ab}\,\omega^b_{(4)c}e^c_{(4)}\ .$$

Again as a pedagogical reminder, we note that were we to extract the coefficient of $e^c_{(4)}e^b_{(4)}$ in the above, we would have to antisymmetrize $\partial_c F_{ab}$ in the indices (cb), but we will not need to do that here. Now to compute $(\omega^2)^\theta{}_a = \omega^\theta{}_m\omega^m{}_a = \omega^\theta{}_b\omega^b{}_a + 0$, where we have used $\omega^\theta{}_\theta = 0$.

When simplifying this expression, it is important to keep track of whether a particular term is a form or a number. In particular, F_{bc} is a number, $\omega^b_{(4)a}$ is a 1-form, and $e^c_{(4)}$ is a 1-form. Therefore:

$$(\omega^2)^\theta{}_a = (\frac{1}{2}a\,F_{bc}e^c_{(4)})(\omega^b_{(4)a} - \frac{1}{2}a\,\eta_{\theta\theta}F^b{}_a e^\theta)$$

$$= -\frac{1}{2}a\,F_{bc}\,\omega^b_{(4)a}e^c_{(4)} - \frac{1}{4}a^2\,\eta_{\theta\theta}F_{bc}F^b{}_a e^c_{(4)}e^\theta$$

where in the first term, the minus sign was generated from moving the 1-form $e^c_{(4)}$ to the right of the 1-form $\omega^b_{(4)a}$. In contrast, no minus sign was generated in the second term when moving the 1-form $e^c_{(4)}$ to the right of the number $F^b{}_a$. Putting it all together, we have[40]

$$R^\theta{}_a = \frac{1}{2}a\,\partial_c F_{ab}\,e^c_{(4)}e^b_{(4)} - \frac{1}{2}a\,F_{ab}\,\omega^b_{(4)c}e^c_{(4)} - \frac{1}{2}a\,F_{bc}\,\omega^b_{(4)a}e^c_{(4)} - \frac{1}{4}a^2\,\eta_{\theta\theta}F_{bc}F^b{}_a\,e^c_{(4)}e^\theta\ .$$

---

[40]Let us again pause to check the dimensions. F_{ab} has dimensions of $1/\text{length}^2$, and $e^c_{(4)}$ has dimensions of length. a has dimensions of length, and ∂_c has dimensions of inverse length. So the first term is dimensionless. Since $\omega^b_{(4)c}$ is dimensionless, the second and third terms are also dimensionless. The fourth term has two powers of length from a^2, one power of length from each of $e^c_{(4)}$ and e^θ, which are indeed canceled by four powers of inverse length from the product $F_{bc}F^b{}_a$.

To get the curvature scalar, we need only the components $R^m{}_{amc} = R^b{}_{abc} + R^\theta{}_{a\theta c}$, so we need to extract $R^\theta{}_{a\theta b}$ from the above expression for $R^\theta{}_a$. Since $R^\theta{}_a = \frac{1}{2}R^\theta{}_{amn}e^m e^n = \frac{1}{2}R^\theta{}_{abc}e^b_{(4)}e^c_{(4)} + 2 \times \frac{1}{2}R^\theta{}_{a\theta c}e^\theta e^c_{(4)}$, we compare "$R^\theta{}_a = -\frac{1}{4}\eta_{\theta\theta}F_{bc}F^b{}_a e^c_{(4)}e^\theta + ... = +\frac{1}{4}\eta_{\theta\theta}F_{bc}F^b{}_a e^\theta e^c_{(4)} + ...$" with "$R^\theta{}_a = R^\theta{}_{a\theta c}e^\theta e^c_{(4)} + ...$" to get

$$R^\theta{}_{a\theta c} = \frac{1}{4}\eta_{\theta\theta}F_{bc}F^b{}_a .$$

Next, recall the previously obtained result

$$R^a{}_{bcd} = (R_{(4)})^a{}_{bcd} - \frac{1}{4}a^2(F^a{}_c F_{bd} - F^a{}_d F_{bc}) - \frac{1}{2}a^2 \, \eta_{\theta\theta}F^a{}_b F_{cd}$$

and contract with $\delta^c{}_a$ to get

$$R^a{}_{bad} = (R_{(4)})^a{}_{bad} - \frac{1}{4}a^2 (0 - F^a{}_d F_{ba}) - \frac{1}{2}a^2 \, \eta_{\theta\theta}F^a{}_b F_{ad}$$
$$= (R_{(4)})^a{}_{bad} + \frac{1}{4}a^2 F^a{}_d(-F_{ab}) - \frac{1}{2}a^2 \, \eta_{\theta\theta}F^a{}_b F_{ad}$$
$$= (R_{(4)})^a{}_{bad} - \frac{1}{4}a^2(1 + 2\eta_{\theta\theta})F_{ab}F^a{}_d .$$

Since we already found $R^\theta{}_{b\theta d} = \frac{1}{4}\eta_{\theta\theta}F_{ab}F^a{}_d$, we arrive at the following expression for $\mathcal{R}_{bd} \equiv R^m{}_{bmd} = R^a{}_{bad} + R^\theta{}_{b\theta d}$:

$$\mathcal{R}_{bd} = (\mathcal{R}_{(4)})_{bd} - \frac{1}{4}a^2(1 + \eta_{\theta\theta})F_{ab}F^a{}_d .$$

We use the calligraphic font to denote $\mathcal{R}_{bd} \equiv R^m{}_{bmd}$ to avoid potential confusion with the previously defined 2-form $R^a{}_c$, which can display a lowered index via $R_{ac} = \eta_{ab}R^b{}_c$. To belabor the point, the object $\mathcal{R}_{ab}$ is a collection of ordinary numbers, while the object R_{ab} is a collection of 2-forms.

To complete the calculation we need the quantity $\mathcal{R}_{\theta\theta} \equiv R^m{}_{\theta m\theta} = R^a{}_{\theta a\theta} = \eta^{ab}R_{b\theta a\theta} = \eta^{ab}R_{\theta b\theta a} = \eta^{ab}\eta_{\theta\theta}R^\theta{}_{b\theta a} = \eta^{ab}\eta_{\theta\theta}(\frac{1}{4}a^2 \, \eta_{\theta\theta}F_{cb}F^c{}_a) = \frac{1}{4}a^2 \, F_{cb}F^{cb}$.

Finally, we can compute the scalar curvature:

$$\mathcal{R} \equiv \eta^{mn}\mathcal{R}_{mn} = \eta^{bd}\mathcal{R}_{bd} + \eta^{\theta\theta}\mathcal{R}_{\theta\theta}$$
$$= \mathcal{R}_{(4)} - \frac{1}{4}a^2(1 + \eta_{\theta\theta})F_{ab}F^{ab} + \eta^{\theta\theta}\frac{1}{4}a^2 \, F_{ab}F^{ab}$$
$$= \mathcal{R}_{(4)} - \frac{1}{4}a^2 \, F_{ab}F^{ab} .$$

Notice that the terms with an explicit factor of $\eta^{\theta\theta} = \eta_{\theta\theta}$ drop out, reflecting the fact that the action should not depend on our convention for the signs in the metric. Switch from flat coordinates to spacetime coordinates by writing $F_{ab} = e^\mu_a e^\nu_b F_{\mu\nu}$ and using $e^\mu_a e^a_\nu = \delta^\mu_\nu$ to get $F_{ab}F^{ab} = F_{\mu\nu}F^{\rho\sigma}e^\mu_a e^\nu_b e^a_\rho e^b_\sigma = F_{\mu\nu}F^{\rho\sigma}\delta^\mu_\rho \delta^\nu_\sigma = F_{\mu\nu}F^{\mu\nu}$. We have therefore obtained the desired result

$$\mathcal{R} = \mathcal{R}_{(4)} - \frac{1}{4}a^2 \, F_{\mu\nu}F^{\mu\nu} .$$

279

6. Generalize the Kaluza-Klein construction by replacing the circles by higher dimensional spheres. Show that Yang-Mills fields emerge.

Solution:

Indices:

$$M = 0, 1, 2, ..., D \leftarrow \text{full spacetime, } \bar{x}^M$$
$$\mu = 0, 1, 2, 3 \leftarrow \text{non-compact part, } x^\mu$$
$$m = 4, 5, ..., D \leftarrow \text{compact part, } \theta^m$$
$$a = 1, 2, ..., K \leftarrow \text{isometry group of compact part; generated by } \xi_a^m(\theta)$$

$$\hat{M} = 0, 1, 2, ..., D \leftarrow \text{locally flat frame of full spacetime}$$
$$\hat{\mu} = 0, 1, 2, 3 \leftarrow \text{locally flat frame of non-compact part}$$
$$\hat{m} = 4, 5, ..., D \leftarrow \text{locally flat frame of compact part}$$

Full metric G_{MN}, full veilbein $E_M^{\hat{M}}$: $G_{MN} = E_M^{\hat{M}} E_N^{\hat{N}} \eta_{\hat{M}\hat{N}}$.

In terms of the compact and non-compact parts:

$$\eta_{\hat{M}\hat{N}} = \begin{pmatrix} \eta_{\mu\nu} & 0 \\ 0 & -\delta_{\hat{m}\hat{n}} \end{pmatrix} , \quad E_M^{\hat{M}}(x, \theta) = \begin{pmatrix} e_\mu^{\hat{\mu}}(x) & -B_\mu^{\hat{m}}(x, \theta) \\ 0 & e_m^{\hat{m}}(\theta) \end{pmatrix} \qquad [B_\mu^{\hat{m}}(x, \theta) \equiv e_m^{\hat{m}}(\theta) B_\mu^m(x, \theta)]$$

$$G_{MN}(x, \theta) = (E\eta E^T)_{MN} = \begin{pmatrix} e_\mu^{\hat{\mu}}(x) & -B_\mu^{\hat{m}}(x, \theta) \\ 0 & e_m^{\hat{m}}(\theta) \end{pmatrix} \begin{pmatrix} \eta_{\hat{\mu}\hat{\nu}} & 0 \\ 0 & -\delta_{\hat{m}\hat{n}} \end{pmatrix} \begin{pmatrix} e_\nu^{\hat{\nu}}(x) & 0 \\ -B_\nu^{\hat{n}}(x, \theta) & e_n^{\hat{n}}(\theta) \end{pmatrix}$$

$$= \begin{pmatrix} e_\mu^{\hat{\mu}}(x)\eta_{\hat{\mu}\hat{\nu}}e_\nu^{\hat{\nu}}(x) - B_\mu^{\hat{m}}(x, \theta)\delta_{\hat{m}\hat{n}}B_\nu^{\hat{n}}(x, \theta) & B_\mu^{\hat{m}}(x, \theta)\delta_{\hat{m}\hat{n}}e_n^{\hat{n}}(\theta) \\ e_m^{\hat{m}}(\theta)\delta_{\hat{m}\hat{n}}B_\nu^{\hat{n}}(x, \theta) & -e_m^{\hat{m}}(\theta)\delta_{\hat{m}\hat{n}}e_n^{\hat{n}}(\theta) \end{pmatrix}$$

$$= \begin{pmatrix} g_{\mu\nu}(x) - B_\mu^m(x, \theta)g_{mn}(\theta)B_\nu^n(x, \theta) & B_\mu^m(x, \theta)g_{mn}(\theta) \\ g_{mn}(\theta)B_\nu^n(x, \theta) & -g_{mn}(\theta) \end{pmatrix}$$

Under the transformation $\bar{x}^M \to \bar{x}'^M$, we have $E_M^{\hat{M}}(\bar{x}) \to E_M'^{\hat{M}}(\bar{x}') = \frac{\partial \bar{x}^N}{\partial \bar{x}'^M} E_N^{\hat{M}}(\bar{x})$. Rearrange the primes:

$$E_M^{\hat{M}}(\bar{x}) = \frac{\partial \bar{x}'^N}{\partial \bar{x}^M} E_N'^{\hat{M}}(\bar{x}') = \frac{\partial x'^\nu}{\partial \bar{x}^M} E_\nu'^{\hat{M}}(\bar{x}') + \frac{\partial \theta'^n}{\partial \bar{x}^M} E_n'^{\hat{M}}(\bar{x}') .$$

Look at $E_\mu^{\hat{m}} = -B_\mu^{\hat{m}}$, with $x'^\mu = x^\mu$ and $\theta'^m = \theta^m + \xi_a^m(\theta)\varepsilon^a(x)$:

$$-B_\mu^{\hat{m}}(x, \theta) = -\frac{\partial x'^\nu}{\partial x^\mu} B_\nu'^{\hat{m}}(x', \theta') + \frac{\partial \theta'^n}{\partial x^\mu} e_n'^{\hat{m}}(\theta')$$
$$= -B_\mu'^{\hat{m}}(x, \theta') + \xi_a^n(\theta)\partial_\mu\varepsilon^a(x)e_n'^{\hat{m}}(\theta') .$$

Write $B_\mu^m(x, \theta) \equiv \xi_a^m(\theta) A_\mu^a(x)$ to get:

$$\xi_a^{\hat{m}}(\theta) A_\mu^a(x) = \xi_a'^{\hat{m}}(\theta') A_\mu'^a(x) - \xi_a^n(\theta) \partial_\mu \varepsilon^a(x) e_n'^{\hat{m}}(\theta')$$

Be mindful of the primes on $\xi_a'^{\hat{m}}(\theta')$ on the left-hand side: we have to ask how a Killing vector transforms under an infinitesimal transformation of the compact coordinates. For $\theta^m \to \theta'^m = (\Lambda^{-1})^m{}_n \theta^n$, with $(\Lambda^{-1})^m{}_n = \frac{\partial \theta'^m}{\partial \theta^n} = \delta_n^m + \partial_n \xi_a^m(\theta) \varepsilon^a(x)$, we have:

$$\begin{aligned}
\xi_a^n(\theta) \to \xi_a'^n(\theta') &= \Lambda^n{}_p \xi_a^p(\Lambda^{-1}\theta) \\
&= \left[\delta_p^n - \partial_p \xi_b^n(\theta) \varepsilon^b(x)\right] \left[\xi_a^p(\theta) + \xi_b^q(\theta)\varepsilon^b(x)\partial_q\xi_a^p(\theta)\right] + O(\varepsilon^2) \\
&= \xi_a^n(\theta) - \left[\xi_a(\theta)\cdot\partial\xi_b^n(\theta) - \xi_b(\theta)\cdot\partial\xi_a^p(\theta)\right]\varepsilon^b(x) \\
&= \xi_a^n(\theta) + g f_{ab}{}^c \xi_c^n(\theta) \varepsilon^b(x)
\end{aligned}$$

where in the last equality we have used Killing's equation. With this and with $e_m^{\hat{m}}(\theta) = e_m'^{\hat{m}}(\theta')$ [since $\frac{x'^\nu}{\partial\theta^m} = 0$], we can strip off the overall contraction with $\xi_a^{\hat{m}}(\theta) \equiv e_m^{\hat{m}}(\theta)\xi_a^m(\theta)$ to obtain

$$A_\mu^a(x) = A_\mu'^a(x) + g f_{bc}{}^a A_\mu'^b(x)\varepsilon^c(x) - \partial_\mu \varepsilon^a(x) .$$

For further discussion, see A. Salam, J. Strathdee, "On Kaluza-Klein theory," Ann. Phys. 141, 316-352 (1982) and F. Cianfrani, G. Montani, "Non Abelian gauge symmetries induced by the unobservability of extra-dimensions in a Kaluza-Klein approach," Mod.Phys.Lett. A21 (2006) 265-274 (arXiv:gr-qc/0511100v1).

8. The veilbeins for a spacetime with Minkowski metric is defined by $g_{\mu\nu}(x) = e_\mu^a(x)\eta_{ab}e_\nu^b(x)$, where the Minkowski metric η_{ab} replaces the Euclidean metric δ_{ab}. The indices a and b are to be contracted with η_{ab}. For example, $R^{ab} = d\omega^{ab} + \omega^{ac}\eta_{cd}\omega^{db}$. Show that everything goes through as expected.

Solution:

We will follow the discussion beginning on p. 443 replacing δ_{ab} with η_{ab}. We begin with

$$g_{\mu\nu}(x) = e_\mu^a(x)\eta_{ab}e_\nu^b(x) \qquad (16_M)$$

For your convenience, we will label the equations as in the chapter, with the subscript M to denote "Minkowski." Consider the "Minkowskian 2-sphere" defined by the line element

$$ds^2 = dt^2 - \sin^2 t\, d\varphi^2 \ .$$

From the metric ($g_{tt} = 1$, $g_{\varphi\varphi} = \sin^2 t$) we can read off $e_t^1 = 1$ and $e_\varphi^2 = \sin t$, with all other components zero. (Remember that $\eta_{ab} = \text{diag}(+1, -1)$ in this case!) We define the 1-forms $e^a = e_\mu^a dx^\mu$, so that $e^1 = dt$, $e^2 = \sin t\, d\varphi$. Now define the curvature 1-form

$$de^a = -\omega^a{}_b e^b \qquad (17_M)$$

where now we are careful to use upper and lower indices because we have a nontrivial norm on the spacetime. In our baby example, we have $de^1 = 0$ and $de^2 = \cos t\, dt\, d\varphi$, so as in the text the connection has only one nonvanishing component $\omega^{12} = -\omega^{21} = -\cos t\, d\varphi$.

In the Euclidean case, the rotation $e_\mu^a(x) = O^a{}_b(x)e_\mu'^b(x)$ leaves the metric invariant, that is $g_{\mu\nu}(x) = e_\mu^a(x)\delta_{ab}e_\nu^b(x) = e_\mu'^a(x)\delta_{ab}e_\nu'^b(x)$ if $O^T O = 1$. This time, under the same transformation $e_\mu^a(x) = O^a{}_b(x)e_\mu'^b(x)$, we have (suppressing the spacetime arguments)

$$g_{\mu\nu} = e_\mu^a\eta_{ab}e_\nu^b = O^a{}_c e_\mu^c\eta_{ab}O^b{}_d e_\nu^d = e_\mu^c(O^T){}_c{}^a\eta_{ab}O^b{}_d e_\nu^d = e_\mu^c\eta_{cd}e_\nu^d$$

only when $O^T\eta O = \eta$. This is precisely the definition of the Lorentz transformation. That is, the rotation $e_\mu^a = O^a{}_b e^b$ leaves the curved metric $g_{\mu\nu}$ invariant if O leaves the flat metric η_{ab} invariant. Everything else up through equation (20) on p. 444 works out in a formally identical manner, including the curvature of $+1$ in our Minkowskian 2-sphere.

VIII.3 Effective Field Theory

1. Consider

$$\mathcal{L} = \frac{1}{2}\left[(\partial\varphi_1)^2 + (\partial\varphi_2)^2\right] - \lambda(\varphi_1^4 + \varphi_2^4) - g\,\varphi_1^2\varphi_2^2$$

We have taken the $O(2)$ theory from chapter I.10 and broken the symmetry explicitly. Work out the renormalization group flow in the (λg)-plane and draw your own conclusions.

Solution:

First let us make a few changes in notation. Normalize the couplings such that we won't have to worry about unnecessary numerical factors in the Feynman rules, meaning $\lambda \to \frac{1}{4!}\lambda$ and $g \to \frac{1}{2^2}g$. To eliminate notational clutter, write $\varphi_1 \equiv \varphi$ and $\chi \equiv \varphi_2$. For this problem we will use dimensional regularization and therefore separate the mass parameter $\tilde{\mu}$ from the couplings: $\lambda \to \lambda\tilde{\mu}^\varepsilon$ and $g \to g\tilde{\mu}^\varepsilon$. The Lagrangian is

$$\mathcal{L} = \frac{1}{2}Z\left[(\partial\varphi)^2 + (\partial\chi)^2\right] - \frac{1}{24}Z_\lambda\lambda\tilde{\mu}^\varepsilon(\varphi^4 + \chi^4) - \frac{1}{4}Z_g g\tilde{\mu}^\varepsilon\,\varphi^2\chi^2$$

$$= \frac{1}{2}\left[(\partial\varphi_0)^2 + (\partial\chi_0)^2\right] - \frac{1}{24}\lambda_0(\varphi_0^4 + \chi_0^4) - \frac{1}{4}g_0\,\varphi_0^2\chi_0^2$$

In the first line we have included the renormalizing Z-factors, or equivalently the counterterms $A_i \equiv Z_i - 1$. In the second line we have written the Lagrangian in terms of bare fields and parameters, denoted by the subscript 0. For future reference, we will need the relationships between the bare parameters and the renormalized parameters:

$$\lambda_0 = Z^{-2}Z_\lambda\lambda\tilde{\mu}^\varepsilon \quad \text{and} \quad g_0 = Z^{-2}Z_g g\tilde{\mu}^\varepsilon$$

The Feynman rules are as follows. We use a solid line for the φ propagator, and we use a dashed line for the χ propagator. Since we assume these fields have the same mass (namely $m^2 = 0$), both propagators are equal, so that (dashed line) = (solid line) = $i\Delta(k)$, where $\Delta(k) \equiv 1/(k^2 - m^2) = 1/k^2$. A vertex connecting 4 solid lines equals a vertex connecting 4 dashed lines, both of which equal $-iZ_\lambda\lambda\tilde{\mu}^\varepsilon$. A vertex connecting 2 solid lines and 2 dashed lines equals $-iZ_g g\tilde{\mu}^\varepsilon$.

At 1-loop order, the propagators do not get renormalized, so $Z = 1$ at 1-loop. The φ^4 vertex has the following 1-loop corrections:

283

We use conventions for which all momenta flow into the vertex. We also use the notational convention for which the diagram is just the picture as written, and the symmetry factor (in this case $\frac{1}{2}$) is written explicitly multiplying the picture (as shown above). The first 1-loop diagram is:

$$
(1) = (-iZ_\lambda \lambda \tilde{\mu}^\varepsilon)^2 \int \frac{d^d\ell}{(2\pi)^d} i\Delta(\ell) i\Delta(\ell - k_1 - k_4)
$$

$$
= (\lambda \tilde{\mu}^\varepsilon)^2 \int \frac{d^d\ell}{(2\pi)^d} \frac{1}{(\ell^2 - m^2)} \frac{1}{[(\ell - k_1 - k_4)^2 - m^2]}
$$

$$
= (\lambda \tilde{\mu}^\varepsilon)^2 \int \frac{d^d\ell}{(2\pi)^d} \int_0^1 dx \frac{1}{\{x(\ell^2 - m^2) + (1-x)[(\ell - k_1 - k_4)^2 - m^2]\}^2}
$$

The denominator simplifies to $[\ell - (1-x)(k_1 + k_4)]^2 + D$, where $D \equiv x(1-x)(k_1 + k_4)^2 - m^2$. Defining $p \equiv \ell - (1-x)(k_1 + k_4)$, we have

$$
(1) = (\lambda \tilde{\mu}^\varepsilon)^2 \int_0^1 dx \int \frac{d^d p}{(2\pi)^d} \frac{1}{(p^2 + D)^2}
$$

$$
= (\lambda \tilde{\mu}^\varepsilon)^2 \int_0^1 dx \frac{i}{(4\pi)^{d/2}} \Gamma(2 - d/2) D^{-(2-d/2)}
$$

$$
= (\lambda \tilde{\mu}^\varepsilon)^2 \frac{i}{(4\pi)^2} \Gamma\left(\frac{\varepsilon}{2}\right) \int_0^1 dx \left(\frac{4\pi}{D}\right)^{\varepsilon/2}
$$

The other two diagrams with only solid lines are the same as this one but with different dependence on the external momenta. Define $D_i \equiv x(1-x)(k_1 + k_i)^2 - m^2$. Then the three diagrams are

$$
(1) = (\lambda \tilde{\mu}^\varepsilon)^2 \frac{i}{(4\pi)^2} \Gamma\left(\frac{\varepsilon}{2}\right) \int_0^1 dx \left(\frac{4\pi}{D_4}\right)^{\varepsilon/2}
$$

$$
(2) = (\lambda \tilde{\mu}^\varepsilon)^2 \frac{i}{(4\pi)^2} \Gamma\left(\frac{\varepsilon}{2}\right) \int_0^1 dx \left(\frac{4\pi}{D_2}\right)^{\varepsilon/2}
$$

$$
(3) = (\lambda \tilde{\mu}^\varepsilon)^2 \frac{i}{(4\pi)^2} \Gamma\left(\frac{\varepsilon}{2}\right) \int_0^1 dx \left(\frac{4\pi}{D_3}\right)^{\varepsilon/2}
$$

Moreover, the other three diagrams (containing dashed lines) are numerically identical to the three diagrams above, except with the replacement $\lambda \to g$. Therefore, at 1-loop order we have

$$
V_\lambda = -\lambda \tilde{\mu}^\varepsilon \left\{ Z_\lambda - \frac{\lambda}{2(4\pi)^2} \left(1 + \frac{g^2}{\lambda^2}\right) \Gamma\left(\frac{\varepsilon}{2}\right) \int_0^1 dx \left[\left(\frac{4\pi\tilde{\mu}^2}{D_2}\right)^{\varepsilon/2} + \left(\frac{4\pi\tilde{\mu}^2}{D_3}\right)^{\varepsilon/2} + \left(\frac{4\pi\tilde{\mu}^2}{D_4}\right)^{\varepsilon/2} \right] \right\}
$$

Now we need some expansions in $\varepsilon \to 0^+$, namely $C^{\varepsilon/2} = 1 + \frac{\varepsilon}{2} \ln C + O(\varepsilon^2)$ for some number C, and $\Gamma(\varepsilon/2) = \frac{2}{\varepsilon} - \gamma + O(\varepsilon)$, where $\gamma \equiv \int_0^\infty dt\, e^{-t} \ln t \approx 0.577$. Together these imply

$$
\Gamma\left(\frac{\varepsilon}{2}\right) C^{\varepsilon/2} = \frac{2}{\varepsilon} + \ln\left(Ce^{-\gamma}\right) + O(\varepsilon) .
$$

284

Now let $\mu^2 \equiv 4\pi e^{-\gamma} \tilde{\mu}^2$. Using the above expansion and taking $\varepsilon \to 0^+$ wherever possible gives

$$V_\lambda = -\lambda \left\{ Z_\lambda - \frac{\lambda}{2(4\pi)^2} \left(1 + \frac{g^2}{\lambda^2}\right) \left[\frac{6}{\varepsilon} + \int_0^1 dx \left(\ln\left(\frac{\mu^2}{D_2}\right) + \ln\left(\frac{\mu^2}{D_3}\right) + \ln\left(\frac{\mu^2}{D_4}\right)\right)\right] \right\}$$

We choose the "modified minimal subtraction" renormalization scheme, for which the counterterm $A_\lambda = Z_\lambda - 1$ is chosen purely to cancel the $1/\varepsilon$ pole and nothing more. To this order, we therefore have

$$A_\lambda = \frac{3}{(4\pi)^2} \lambda \left(1 + \frac{g^2}{\lambda^2}\right) \frac{1}{\varepsilon} \ .$$

Now we need to renormalize the $\varphi^2 \chi^2$ vertex. At 1-loop, we have

Fortunately, we have already done all of the computational work. The two s-channel diagrams are numerically equal to the s-channel diagram from the φ^4 vertex, except with the replacement $\lambda^2 \to \lambda g$. The t-channel and u-channel diagrams are numerically equal to the t- and u- channel diagrams from the φ^4 vertex, except with the replacement $\lambda^2 \to g^2$. Immediately we get the result

$$V_g = -g \left\{ Z_g - \frac{1}{(4\pi)^2} \left[(\lambda + 2g)\frac{2}{\varepsilon} + (\text{finite})\right] \right\}$$

so that choosing the counterterm $A_g \equiv Z_g - 1$ purely to cancel the divergent piece gives

$$A_g = \frac{2}{(4\pi)^2} (\lambda + 2g) \frac{1}{\varepsilon} \ .$$

Now we compute the beta functions. Define $H(\lambda, g) \equiv \ln(Z^{-2} Z_\lambda)$ and $G(\lambda, g) \equiv \ln(Z^{-2} Z_g)$. Taking the logarithm of the relationships between the bare and renormalized couplings gives

$$\ln \lambda_0 = H(\lambda, g) + \ln \lambda + \varepsilon \ln \tilde{\mu} \quad \text{and} \quad \ln g_0 = G(\lambda, g) + \ln g + \varepsilon \ln \tilde{\mu} \ .$$

The bare parameters are independent of the unphysical parameter μ, so differentiating the above with respect to $\ln \mu$ gives

$$0 = \lambda \frac{dH}{d\ln\mu} + \frac{d\lambda}{d\ln\mu} + \varepsilon\lambda \quad \text{and} \quad 0 = g \frac{dG}{d\ln\mu} + \frac{dg}{d\ln\mu} + \varepsilon g \ .$$

Upon using the chain rule and rearranging a bit, these become

$$0 = \left(\lambda\frac{\partial H}{\partial \lambda} + 1\right)\frac{d\lambda}{d\ln\mu} + \lambda\frac{\partial H}{\partial g}\frac{dg}{d\ln\mu} + \varepsilon\lambda \tag{1}$$

and

$$0 = \left(g\frac{\partial G}{\partial g} + 1\right)\frac{dg}{d\ln\mu} + g\frac{\partial G}{\partial \lambda}\frac{d\lambda}{d\ln\mu} + \varepsilon g \tag{2}$$

The renormalizing Z-factors in the MS-bar renormalization scheme to this order are

$$Z = 0$$

$$Z_\lambda = 1 + \frac{3}{(4\pi)^2}\lambda\left(1 + \frac{g^2}{\lambda^2}\right)\frac{1}{\varepsilon}$$

$$Z_g = 1 + \frac{2}{(4\pi)^2}(\lambda + 2g)\frac{1}{\varepsilon}$$

so the functions H and G are (after expanding $\ln(1 + x) = x + O(x^2)$)

$$H(\lambda, g) = \frac{3}{(4\pi)^2}\lambda\left(1 + \frac{g^2}{\lambda^2}\right)\frac{1}{\varepsilon} \implies \frac{\partial H}{\partial \lambda} = \frac{3}{(4\pi)^2}\left(1 - \frac{g^2}{\lambda^2}\right)\frac{1}{\varepsilon} \text{ and } \frac{\partial H}{\partial g} = \frac{6g}{(4\pi)^2\lambda}\frac{1}{\varepsilon}$$

$$G(\lambda, g) = \frac{2}{(4\pi)^2}(\lambda + 2g)\frac{1}{\varepsilon} \implies \frac{\partial G}{\partial \lambda} = \frac{2}{(4\pi)^2}\frac{1}{\varepsilon} \text{ and } \frac{\partial G}{\partial g} = \frac{4}{(2\pi)^2}\frac{1}{\varepsilon}$$

With these, the renormalization group equations (1) and (2) become

$$0 = \left[\frac{3}{(4\pi)^2}\lambda\left(1 - \frac{g^2}{\lambda^2}\right)\frac{1}{\varepsilon} + 1\right]\frac{d\lambda}{d\ln\mu} + \frac{6}{(4\pi)^2}g\frac{1}{\varepsilon}\frac{dg}{d\ln\mu} + \varepsilon\lambda \tag{1'}$$

$$0 = \left[\frac{4}{(4\pi)^2}g\frac{1}{\varepsilon} + 1\right]\frac{dg}{d\ln\mu} + \frac{2}{(4\pi)^2}g\frac{1}{\varepsilon}\frac{d\lambda}{d\ln\mu} + \varepsilon g \tag{2'}$$

Now as per the usual procedure with dimensional regularization, we write

$$\frac{d\lambda}{d\ln\mu} = -\varepsilon\lambda + \beta_\lambda \text{ and } \frac{dg}{d\ln\mu} = -\varepsilon g + \beta_g$$

and demand that β_λ and β_g be finite in the limit $\varepsilon \to 0^+$. Putting these into equation (1') implies

$$0 = \beta_\lambda - \frac{3}{(4\pi)^2}(\lambda^2 + g^2) + (\text{things we insist sum to zero})$$

while equation (2') implies

$$0 = \beta_g - \frac{2}{(4\pi)^2}(\lambda + 2g)g + (\text{things we insist sum to zero}) .$$

Thus we have the 1-loop beta functions

$$\beta_\lambda = \left(\frac{d\lambda}{d\ln\mu}\right)_{(\text{finite})} = +\frac{3}{(4\pi)^2}(\lambda^2 + g^2) \tag{1''}$$

$$\beta_g = \left(\frac{dg}{d\ln\mu}\right)_{(\text{finite})} = +\frac{2}{(4\pi)^2}(\lambda + 2g)g \tag{2''}$$

286

Before discussing the renormalization group dynamics in the (λ, g)-plane, let us perform a check. There should be a value of the parameter g for which we recover an $O(2)$ symmetric theory. Consider the 2-component vector $\vec{\phi} = (\phi_1, \phi_2)^T$. An $O(2)$-invariant theory would have the interaction Lagrangian

$$\mathcal{L} = -\frac{1}{8}\lambda'(\vec{\phi} \cdot \vec{\phi})^2 = -\frac{1}{8}\lambda' \left(\phi_1^4 + \phi_2^4 + 2\phi_1^2 \phi_2^2 \right)$$

with some coupling λ'. Compare this to our interaction Lagrangian

$$\mathcal{L} = -\frac{1}{4!}\lambda(\varphi^4 + \chi^4) - \frac{1}{4}g\,\varphi^2\chi^2 = -\frac{1}{4!}\left[\lambda(\varphi^4 + \chi^4) + 3!\,g\,\varphi^2\chi^2\right] \ .$$

We recover the $O(2)$ symmetric theory only for the precise value $3!\,g = 2\lambda \implies g = \frac{1}{3}\lambda$. If we plug this particular value into the beta functions, we should recover a single beta function for the coupling λ. If $g = \frac{1}{3}\lambda$, then $g^2 = \frac{1}{9}\lambda^2$, so that $\beta_\lambda = \frac{1}{(4\pi)^2}\frac{10}{3}\lambda^2$. Also, $\beta_g = \frac{2}{(4\pi)^2}(1 + \frac{2}{3})(\frac{1}{3})\lambda^2 = \frac{1}{(4\pi)^2}\frac{10}{9}\lambda^2$. Using the chain rule, we have $dg/d\ln\mu = (dg/d\lambda)(d\lambda/d\ln\mu) = \frac{1}{3}\beta_\lambda$. So $\beta_\lambda = \frac{1}{(4\pi)^2}\frac{10}{3}\lambda^2$, which is exactly what we got before.

Now let us study the dynamics implied by $(1'')$ and $(2'')$. Equation $(1'')$ implies that $d\lambda/d\ln\mu > 0$, so the coupling λ increases in magnitude as the parameter μ increases, irrespective of the signs of λ and g. Equation $(2'')$ is more interesting. Since $d\ln g/d\ln\mu \propto \lambda + 2g$, the relative signs of λ and g change the running of g qualitatively: if $\lambda + 2g < 0$, then the coupling g decreases in magnitude as μ increases. At low energy there is an $O(2)$ symmetric fixed point.

2. Assuming the nonexistence of the right handed neutrino field ν_R (i.e., assuming the minimal particle content of the standard model) write down all $SU(2) \otimes U(1)$ invariant terms that violate lepton number L by 2 and hence construct an effective field theory of the neutrino mass. Of course, by constructing a specific theory one can be much more predictive. Out of the product $l_L l_L$ we can form a Lorentz scalar transforming as either a singlet or triplet under $SU(2)$. Take the singlet case and construct a theory. [Hint: For help, see A. Zee, *Phys. Lett.* 93B: p. 389, 1980.]

Solution:

We use two-component spinor notation for the fermion fields. The neutrinos reside in $SU(2)$ doublets

$$(\ell_a)_i = \begin{pmatrix} \nu_a \\ e_a \end{pmatrix} \sim (2, -\tfrac{1}{2}) \ \text{of} \ SU(2) \otimes U(1)$$

where $i = 1, 2$ labels the fundamental of $SU(2)$, and $a = e, \mu, \tau$ labels the flavor. From this field we can form the $SU(2)$ singlet

$$\ell_a \cdot \ell_b \equiv \varepsilon^{ij}(\ell_a)_i(\ell_b)_j = \nu_a e_b - e_a \nu_b$$

which has hypercharge $2(-\frac{1}{2}) = -1$. Note that this singlet is antisymmetric in the flavor indices. If we introduce a complex scalar field h^+ with hypercharge $+1$ that is a singlet under $SU(2)$, then we can form the interaction terms

$$\mathcal{L}_{h\ell\ell} = \tfrac{1}{2} f_{ab}\, h^+ \ell_a \cdot \ell_b + h.c.$$

where the sums over $a, b = e, \mu, \tau$ are implied, and due to fermi statistics[41] only the antisymmetric part of f contributes.

At this stage, we may simply assign two units of lepton number to the field h^+, and so this interaction does not generate Majorana neutrino masses on its own. Let us now augment the standard model with additional Higgs doublets:

$$(\varphi_I)_i = \begin{pmatrix} \varphi_I^0 \\ \varphi_I^- \end{pmatrix} \sim (2, -\tfrac{1}{2})$$

where again $i = 1, 2$ labels the fundamental of $SU(2)$, while $I = 1, 2, ..., n_H$ labels the flavor of Higgs. This allows for the dimension-3 interaction

$$\mathcal{L}_{h\varphi\varphi} = M_{IJ}\, h^+ \varphi_I \cdot \varphi_J + h.c.$$

where M_{IJ} is an antisymmetric matrix of size $n_H \times n_H$ whose entries have dimensions of mass.

If the Higgs fields give charged lepton masses in the usual way, $\mathcal{L}_{\mathrm{Yuk}} = -y_{ab}^I \varphi_I \cdot \ell_a \bar{e}_b + h.c.$, then they must be assigned lepton number zero. Thus, the clash between $\mathcal{L}_{h\varphi\varphi}$ and $\mathcal{L}_{h\ell\ell}$ breaks lepton number by 2 and thereby generates neutrino masses at 1-loop.

For simplicity, let us assume that only φ_1 gives mass to the charged leptons, and that $n_H = 2$. Then neutrino masses are generated by the diagram
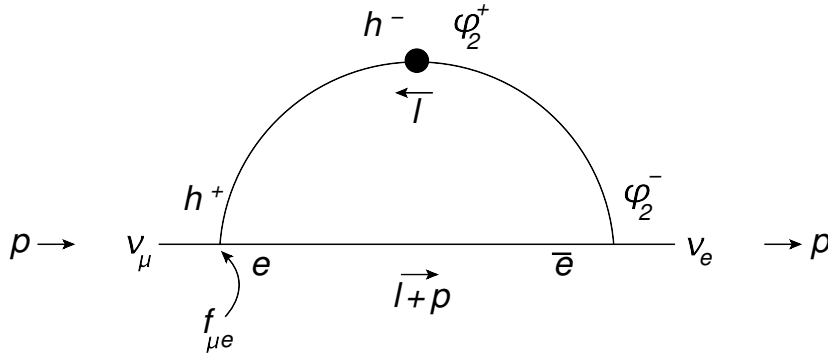
---
[41]Recall that $\nu e = e\nu$ for two Grassmann-valued spinors ν_α and e_α contracted with the antisymmetric tensor $\varepsilon^{\alpha\beta}$. See appendix E.

The diagram is[42]

$$iA_{\mu e}(p) = \int \frac{d^4\ell}{(2\pi)^4} v_e(p)(+iy_e)[iS_e(\ell + p)](+if_{\mu e})u_\mu(p)[i\Delta_{\varphi_2^\pm}(\ell)](+iM_{12}^*\langle\varphi_1^0\rangle^*)[i\Delta_h(\ell)]$$

$$= f_{\mu e}M_{12}^*m_e^2 v_e(p)u_\mu(p)\,\mathcal{I}(p)$$

where

$$\mathcal{I}(p) = -\int \frac{d^4\ell}{(2\pi)^4} \frac{1}{(\ell + p)^2 - m_e^2} \frac{1}{\ell^2 - M_{\varphi_2^\pm}^2} \frac{1}{\ell^2 - M_h^2}\ .$$

We want the mass term, so take $p \to 0$. Also, $m_e \ll M_{\varphi_2^\pm}$, M_h so take $m_e \to 0$ inside the integral. Define $\mathcal{I}_0 \equiv \lim_{m_e \to 0} \mathcal{I}(p = 0)$. Let $M_>$ denote the greater of $M_{\varphi_2^\pm}$ and M_h, and $M_<$ the lesser of the two.

$$\mathcal{I}_0 = -\int \frac{d^4\ell}{(2\pi)^4} \frac{1}{\ell^2} \frac{1}{\ell^2 - M_>^2} \frac{1}{\ell^2 - M_<^2}$$

$$= -\Gamma(3) \int_0^1 dx_1\, dx_2\, dx_3\, \delta(\sum_{i=1}^3 x_i - 1) \int \frac{d^4\ell}{(2\pi)^4} \frac{1}{[x_1\ell^2 + x_2(\ell^2 - M_>^2) + x_3(\ell^2 - M_<^2)]^3}$$

$$= -\Gamma(3) \int_0^1 dx_2 \int_0^{1-x_2} dx_3 \int \frac{d^4\ell}{(2\pi)^4} \frac{1}{(\ell^2 - D)^3}\ , \quad D = x_2 M_>^2 + x_3 M_<^2$$

$$= +\Gamma(3) \int_0^1 dx_2 \int_0^{1-x_2} dx_3\, \frac{i}{(4\pi)^2} \frac{1}{\Gamma(3)} \frac{1}{D}$$

$$= \frac{i}{(4\pi)^2} \frac{1}{M_>^2} \int_0^1 dx_2 \int_0^{1-x_2} dx_3 \frac{1}{x_2 + rx_3}\ , \quad r \equiv \frac{M_<^2}{M_>^2} < 1$$

$$\approx \frac{i}{(4\pi)^2} \frac{1}{M_>^2} \ln\left(\frac{M_>^2}{M_<^2}\right)$$

where we have taken $M_>^2 \gg M_<^2$. We have

$$(m_\nu)_{e\mu} = A_{\mu e}(0) + A_{e\mu}(0) \approx f_{e\mu}(m_\mu^2 - m_e^2)M_{12}^* \frac{1}{(4\pi)^2} \frac{1}{M_>^2} \ln\left(\frac{M_>^2}{M_<^2}\right)$$

where we have used $f_{\mu e} = -f_{e\mu}$.

---

[42]We are using the two-component spinor notation, in which the Fourier expansion of the free neutrino fields is

$$\nu(x) = \sum_s \int (dp)[b_s(\vec{p})u_s(\vec{p})\, e^{-ip\cdot x} + d_s^\dagger(\vec{p})v_s(\vec{p})\, e^{ip\cdot x}]\ .$$

Therefore:

$$\int d^4x\, \nu_e(x)\nu_\mu(x) = \sum_{s_e,s_\mu} \int (dp_e) \int (dp_\mu)(2\pi)^4\delta^4(p_\mu - p_e)\, v_{s_e}(\vec{p}_e)\, u_{s_\mu}(\vec{p}_\mu)d_{s_e}^\dagger(\vec{p}_e)\, b_{s_\mu}(\vec{p}_\mu) + ...\ .$$

So we want the coefficient of $v_{s_e}(\vec{p}_e)\, u_{s_\mu}(\vec{p}_\mu)$ in the 1-loop diagram.

The distinctive prediction of this model is that the diagonal entries of m_ν are zero in the basis for which the charged lepton mass matrix is diagonal. Present neutrino oscillation data rule out such a mass matrix in the absence of further structure.

See the reference along with D. Chang and A. Zee, "Radiatively Induced Neutrino Majorana Masses and Oscillation," arXiv:hep-ph/9912380v1 and R. A. Porto and A. Zee, "Neutrino Mixing and the Private Higgs," arXiv:0807.0612v1 [hep-ph] for further discussion.

3. Let A, B, C, D denote four spin-1/2 fields and label their handedness by a subscript: $\gamma^5 A_h = h A_h$ with $h = \pm 1$. Thus, A_+ is right handed, A_- is left handed, and so on. Show that

$$(A_h B_h)(C_{-h} D_{-h}) = -\frac{1}{2}(A_h \gamma^\mu D_{-h})(C_{-h} \gamma_\mu B_h)$$

This is an example of a broad class of identities known as Fierz identities (some of which we will need in discussing supersymmetry.) Argue that if proton decay proceeds in lowest order from the exchange of a vector particle then only the terms $(\tilde{l}_L C q_L)(u_R C d_R)$ and $(e_R C u_R)(\tilde{q}_L C q_L)$ are allowed in the Lagrangian.

Solution:

In 2-component ("Weyl") notation, we have

$$\sigma^\mu_{\alpha\dot\alpha} \bar\sigma_\mu^{\dot\beta\beta} = 2\, \delta_\alpha{}^\beta \delta_{\dot\alpha}{}^{\dot\beta}$$

Therefore:

$$
\begin{aligned}
(a\sigma^\mu d^\dagger)(c^\dagger \bar\sigma_\mu b) &= 2\, a^\alpha (\sigma^\mu)_{\alpha\dot\alpha} d^{\dagger\dot\alpha} c^\dagger_{\dot\beta} \bar\sigma_\mu^{\dot\beta\beta} b_\beta \\
&= 2\, a^\alpha d^{\dagger\dot\alpha} \delta_\alpha{}^\beta \delta_{\dot\alpha}{}^{\dot\beta} c^\dagger_{\dot\beta} b_\beta \\
&= 2\, a^\alpha d^{\dagger\dot\alpha} c^\dagger_{\dot\alpha} b_\alpha \\
&= (-)2\, a^\alpha b_\alpha\, d^\dagger_{\dot\alpha} c^{\dagger\dot\alpha} \\
&= -2(ab)(d^\dagger c^\dagger) \\
&= -2(ab)(c^\dagger d^\dagger)
\end{aligned}
$$

So

$$(ab)(c^\dagger d^\dagger) = -\frac{1}{2}(a\sigma^\mu d^\dagger)(c^\dagger \bar\sigma_\mu b)$$

Put this in Dirac notation

$$A_L = \begin{pmatrix} a \\ 0 \end{pmatrix}, \quad B_L = \begin{pmatrix} b \\ 0 \end{pmatrix}, \quad C_R = \begin{pmatrix} 0 \\ c^\dagger \end{pmatrix}, \quad D_R = \begin{pmatrix} 0 \\ d^\dagger \end{pmatrix}$$

to get

$$(A_L B_L)(C_R D_R) = -\frac{1}{2}(A_L \gamma^\mu D_R)(C_R \gamma_\mu B_L)$$

290

As for proton decay, we have already seen in problem VII.6.2 that the low-energy effective theory derived from $SU(5)$ unification contains vector bosons with electric charges $\frac{4}{3}$ and $\frac{1}{3}$ that are triplets under $SU(3)$ and doublets under $SU(2)$. These couple to the currents

$$J_\alpha^i = \bar{d}_\alpha \sigma^\mu \ell^{\dagger i} - \varepsilon_{\alpha\beta\gamma} \bar{u}^{\dagger\gamma} \bar{\sigma}^\mu \varepsilon^{ij} q_j^\beta + q_\alpha^{\dagger i} \bar{\sigma}^\mu \bar{e}$$

$$= \overline{D_{R\alpha}} \gamma^\mu L_R^c - \varepsilon_{\alpha\beta\gamma} \overline{U_L^{c\gamma}} \gamma^\mu \varepsilon^{ij} Q_{Lj}^\beta + \overline{Q_{L\alpha}^i} \gamma^\mu E_L^c$$

In the first line we have written the current in 2-component notation, and in the second line we have used Dirac notation. Here $i = 1, 2$ denotes the fundamental of $SU(2)$, and $\alpha = 1, 2, 3$ denotes the fundamental or anti-fundamental of $SU(3)$ depending on the index height.

Integrating out the vector bosons that couple to these currents results in the effective Lagrangian $J^\dagger J$. The operators of the form $qq q\ell$ contained in this Lagrangian are:[43]

$$\varepsilon_{\alpha\beta\gamma}(\bar{u}^\dagger \bar{\sigma}^\mu \varepsilon^{ij} q_j^\beta)(\ell_i \sigma_\mu \bar{d}^{\dagger\alpha}) = -2\varepsilon_{\alpha\beta\gamma} \varepsilon^{ij}(\ell_i q_j^\beta)(\bar{u}^{\dagger\gamma} \bar{d}^{\dagger\alpha}) = -2\varepsilon_{\alpha\beta\gamma}(\nu d^\beta - e u^\beta)(\bar{u}^{\dagger\gamma} \bar{d}^{\dagger\alpha})$$

$$= -2\varepsilon_{\alpha\beta\gamma}(\overline{N_R^c} D_L^\beta - \overline{E_R^c} U_L^\beta)(\overline{U_L^{c\gamma}} D_R^\alpha)$$

$$\varepsilon_{\alpha\beta\gamma}(\bar{u}^{\dagger\gamma} \bar{\sigma}^\mu \varepsilon^{ij} q_j^\beta)(\bar{e}^\dagger \bar{\sigma}_\mu q_i^\alpha) = +2\varepsilon_{\alpha\beta\gamma} \varepsilon^{ij}(q_i^\alpha q_j^\beta)(\bar{u}^{\dagger\gamma} \bar{e}^\dagger) = +4\varepsilon_{\alpha\beta\gamma}(u^\alpha d^\beta)(\bar{u}^{\dagger\gamma} \bar{e}^\dagger)$$

$$= +4\varepsilon_{\alpha\beta\gamma}(\overline{U_R^{c\alpha}} D_L^\beta)(\overline{U_L^{c\gamma}} E_R)$$

These are the first two effective operators given at the top of p. 456.

The other possibility consistent with $SU(3) \otimes SU(2) \otimes U(1)$ invariance and ordinary matter fields is that proton decay is mediated through vector bosons with electric charges $\frac{2}{3}$ and $-\frac{1}{3}$. These would couple to the current

$$J'^i_\alpha = \bar{u}_\alpha \sigma^\mu \ell^{\dagger i} - \varepsilon_{\alpha\beta\gamma} \bar{d}^{\dagger\gamma} \bar{\sigma}^\mu \varepsilon^{ij} q_j^\beta$$

$$= \overline{U_{R\alpha}} \gamma^\mu L_R^c - \varepsilon_{\alpha\beta\gamma} \overline{D_L^{c\gamma}} \gamma^\mu \varepsilon^{ij} Q_{Lj}^\beta \ .$$

The $i = 1$ component has electric charge $Q(\bar{u}\nu^\dagger) = -Q(u) - Q(\nu) = -\frac{2}{3}$, and the $i = 2$ component has electric charge $Q(\bar{u}e^\dagger) = -Q(u) - Q(e) = -\frac{2}{3} - (-1) = +\frac{1}{3}$. The operator contained in $J'^\dagger J'$ relevant for proton decay is:

$$\varepsilon^{\alpha\beta\gamma}(\varepsilon_{ij} q_\beta^{\dagger j} \bar{\sigma}^\mu \bar{d}_\gamma)(\bar{u}_\alpha \sigma_\mu \ell^{\dagger i}) = -2\varepsilon^{\alpha\beta\gamma} \varepsilon_{ij}(\bar{u}_\alpha \bar{d}_\gamma)(q_\beta^{\dagger j} \ell^{\dagger i})$$

which is the hermitian conjugate of the first operator from $J^\dagger J$.

---

[43]To translate between 2-component spinors and Dirac spinors, consider the following example: $E = \begin{pmatrix} e \\ \bar{e}^\dagger \end{pmatrix} \implies E^c = \begin{pmatrix} \bar{e} \\ e^\dagger \end{pmatrix} \implies E_R^c = \begin{pmatrix} 0 \\ e^\dagger \end{pmatrix} \implies \overline{E_R^c} = (e, 0)$. Then $U_L = \begin{pmatrix} u \\ 0 \end{pmatrix} \implies \overline{E_R^c} U_L = eu$.

4. Given the conclusion of the previous exercise show that the decay rate for the processes $p \to \pi^+ + \bar{\nu}$, $p \to \pi^0 + e^+$, $n \to \pi^0 + \bar{\nu}$, and $n \to \pi^- + e^+$ are proportional to each other, with the proportionality factors determined by a single unknown constant [the ratio of the coefficients of $(\bar{l}_L C q_L)(u_R C d_R)$ and $(e_R C u_R)(\tilde{q}_L C q_L)$].

For help on these last three exercises see S. Weinberg, *Phys. Rev. Lett.* 43:1566, 1979; F. Wilczek and A. Zee, ibid. p. 1571; H. A. Weldon and A. Zee, *Nucl. Phys.* B173: 269, 1980.

Solution:

The hadronic part of the first operator transforms as a doublet under isospin. Since $\tilde{q}_L C q_L = 2\tilde{d}_L C u_L$ and parity acts on a Dirac spinor ψ as $\psi_L \to i\gamma^0 \psi_R$, the hadronic part of the second operator is just the parity transform of the first component of the hadronic part of the first operator. Let the coefficients of the two operators be $C_{1,2}$, respectively, let us denote their ratio by $C_2/C_1 = r$. Then the transformation properties under isospin imply the relations

$$\Gamma(p \to \pi^0 e^+) = \tfrac{1}{2}\Gamma(n \to \pi^- e^+) = \tfrac{1}{2}(1 + r^2)\Gamma(p \to \pi^+ \bar{\nu}) = (1 + r^2)\Gamma(n \to \pi^0 \bar{\nu}) \ .$$

As we saw in problem VII.6.2, at tree level in $SU(5)$ unification we have $r = 2$. For a discussion of renormalization group corrections to r, see F. Wilczek and A. Zee, "Operator analysis of nucleon decay," Phys. Rev. Lett. Vol. 43 No. 21, 19 Nov 1979.

5. Imagine a mythical (and presumably impossible) race of physicists who only understand physics at energies less than the electron mass m_e. They manage to write down the effective field theory for the one particle they know, the photon,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{m_e^4}[a(F_{\mu\nu}F^{\mu\nu})^2 + b(F_{\mu\nu}\tilde{F}^{\mu\nu})^2] + ...$$

with $\tilde{F}^{\mu\nu} = \tfrac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$ the dual field strength as usual and a and b two dimensionless constants presumably of order unity.

a) Show that $\mathcal{L}$ respects charge conjugation ($A \to -A$ in this context), parity, and time reversal, (and of course gauge invariance).

Solution:

Since $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, each term in the Lagrangian has two powers of A and is thereby invariant under $A \to -A$. The term $F_{\mu\nu}F^{\mu\nu}$ is invariant under parity and time reversal, as usual. The term $F_{\mu\nu}\tilde{F}^{\mu\nu} = \tfrac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$ involves the antisymmetric tensor $\varepsilon^{\mu\nu\rho\sigma}$ and therefore breaks parity and time reversal. However, it appears squared in the effective Lagrangian depicted above, so the Lagrangian preserves parity and time reversal. Since $F_{\mu\nu}$ is gauge invariant, the Lagrangian is also gauge invariant.

b) Draw the Feynman diagrams that give rise to the two dimension 8 terms shown. The coefficients a and b were calculated by Euler and Kockel in 1935 and by Heisenberg and Euler in 1936, quite a feat since they did not know about Feynman diagrams and any of the modern quantum field theory setup.

Solution:

The electron propagator at zero momentum is $\frac{1}{m_e}$, so we will need four electron propagators to generate the dimension-8 terms. We can draw the "box" diagram with four external photons:
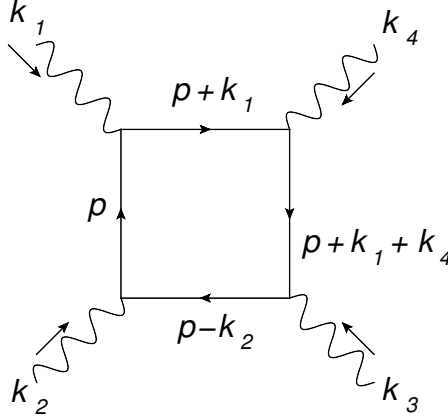
This is similar to the diagram in problem III.8.2, except now the internal lines are fermion propagators and the external lines are spin-1 bosons rather than spin-0. To this diagram (call it $\mathcal{M}_{1234}$) we add the crossed diagrams:

$$\mathcal{M} = \mathcal{M}_{1234} + \mathcal{M}_{1243} + \mathcal{M}_{1324} \ .$$

Recall also that the closed fermion loop generates a factor of (-1) in the amplitude.

Applying the momentum-space Feynman rules, the diagram $\mathcal{M}_{1234}$ is

$$\mathcal{M}_{1234} = \frac{e^4}{(2\pi)^4} \int_\Lambda d^4 p \frac{\mathrm{tr}[\not{\epsilon}_1 (\not{p} + m) \not{\epsilon}_2 (\not{p} - \not{k}_2 + m) \not{\epsilon}_3 (\not{p} + \not{k}_1 + \not{k}_4 + m) \not{\epsilon}_4 (\not{p} + \not{k}_1 + m)]}{[p^2 - m^2][(p - k_2)^2 - m^2][(p + k_1 + k_4)^2 - m^2][(p + k_1)^2 - m^2]}$$

where $m \equiv m_e$ is the mass of the electron, and the integral is cut off at an arbitrary momentum scale Λ, which is of order m. At energies much less than m, we can contract the box to a point interaction to get the two dimension-8 terms $(F_{\mu\nu} F^{\mu\nu})^2$ and[44] $(F_{\mu\nu} \tilde{F}^{\mu\nu})^2$.

---

[44]Since the Lagrangian is invariant under parity and time reversal, it should be possible to write it without the $\varepsilon^{\mu\nu\rho\sigma}$ symbol present in $\tilde{F}^{\mu\nu}$. Indeed, using the identity

$$\varepsilon^{\mu_1...\mu_4} \varepsilon_{\nu_1...\nu_4} = 4! \, \delta^{\mu_1}_{[\nu_1}...\delta^{\mu_4}_{\nu_4]}$$

we have

$$(F_{\mu\nu} \tilde{F}^{\mu\nu})^2 = F_{\mu\nu} F^{\nu\rho} F_{\rho\sigma} F^{\sigma\mu}$$

which is manifestly independent of $\varepsilon^{\mu\nu\rho\sigma}$.

In momentum space, this corresponds to an expansion in powers of k_i/m.

The momentum-space electromagnetic field $F \sim \partial A$ is of the form $F \sim k\varepsilon$, so that the terms $(F^2)^2$ and $(F\tilde{F})^2$ are schematically of the form $kkkk\,\varepsilon\varepsilon\varepsilon\varepsilon$. Thus we are to compute the amplitude $\mathcal{M}$ and extract from it all terms of fourth order in external momenta.

After an involved calculation, one finds[45]

$$a = \frac{1}{4}\,x \quad \text{and} \quad b = \frac{7}{16}\,x \quad \text{where} \quad x \equiv \frac{e^4}{360\pi^2} \ .$$

c) Explain why dimension 6 terms are absent in $\mathcal{L}$. [Hint: One possible term is $\partial_\lambda F_{\mu\nu}\partial^\lambda F^{\mu\nu}$.]

Solution:

Dimension-6 terms can be removed by a field redefinition. See the appendix on p. 458.

d) Our mythical physicists do not know about the electron, but they are getting excited. They are going to start doing photon-photon scattering experiments with a machine called LPC that could produce photons with energy greater than m_e. Discuss what they will see. Apply unitarity and the Cutkosky rules.

Solution:

If we cut the box diagram down the middle we find a process in which two photons annihilate into an electron-positron pair. When the center-of-mass energy reaches the threshold $\sqrt{s} = 2m_e$, then the process $\gamma\gamma \to e^+e^-$ will occur. See problem III.8.3 for discussion of a similar process.

---

[45]arXiv:hep-th/0406216v1

6. Use the effective field theory approach to show that the scattering cross section of light on an electrically neutral spin $1/2$ particle (such as the neutron) goes like $\sigma \propto \omega^2$ to leading order, not ω^4. Argue further that the constant of proportionality can be fixed in terms of the magnetic moment μ of the particle. [Historical note: This result was first obtained in 1954 by F. Low (*Phys. Rev.* 96: 1428) and by M. Gell-Mann and Murph L. Goldberger (*Phys. Rev.* 96: 1433) using much more elaborate arguments.]

Solution:

The operator of lowest dimension we can write down is $F_{\mu\nu}\bar{\psi}\sigma^{\mu\nu}\psi$, which is dimension $[F] + 2[\psi] = 2 + 2(3/2) = 5$. This has only one factor of $F_{\mu\nu}$, meaning one derivative on A_μ. The amplitude gets one power of frequency ω, so the cross section goes like ω^2.

Since the spin operator is $S_z = \frac{1}{2}\sigma^{12} = \frac{1}{2}\begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}$ and a magnetic field pointing in the z-direction is $B = \frac{1}{2}\varepsilon^{3ij}F_{ij} = F_{12}$, the above operator is $B\bar{\psi}S_z\psi$. The magnetic moment μ is defined by the response of the spin to an external magnetic field, or in other words $\mu = \langle e|\bar{\psi}S_z\psi|e\rangle$, where $|e\rangle$ is a state with one electron.

VIII.4 Supersymmetry

1. Construct the Wess-Zumino Lagrangian by the trial and error approach.

Solution:

The goal is to construct the Lagrangian in equation (12) on p. 467, namely

$$\mathcal{L} = \partial\varphi^\dagger\partial\varphi + i\psi^\dagger\bar{\sigma}^\mu\partial_\mu\psi + F^\dagger F - \left[m\left(F\varphi - \frac{1}{2}\psi\psi\right) + g\varphi\left(F\varphi - \psi\psi\right) + h.c.\right].$$

With the benefit of hindsight, let us write only the terms we know we need: $\mathcal{L} = \mathcal{L}_\varphi + \mathcal{L}_\psi + \mathcal{L}_{\varphi\psi} + \mathcal{L}_F$, where

$$\mathcal{L}_\varphi = \partial_\mu\varphi^\dagger\partial^\mu\varphi - (\frac{1}{2}\mu^2\varphi^2 + h.c.)$$
$$\mathcal{L}_\psi = \frac{1}{2}i\psi^\dagger\bar{\sigma}^\mu\partial_\mu\psi - \frac{1}{2}m\,\psi\psi + h.c.$$
$$\mathcal{L}_{\varphi\psi} = g\,\varphi\psi\psi + h.c.$$
$$\mathcal{L}_F = F^\dagger F - (\kappa\,F\varphi + \lambda\,F\varphi^2 + h.c.)$$

We will for simplicity also assume that the couplings are all real, even though this is not necessarily true in general. (Note that we have written $\frac{1}{2}i\psi^\dagger\bar{\sigma}^\mu\partial_\mu\psi + h.c. = i\psi^\dagger\bar{\sigma}^\mu\partial_\mu\psi$ up to a total derivative.)

The infinitesimal transformations have the form:

$$\delta\varphi = \xi^\alpha \psi_\alpha \ , \quad \delta\psi_\alpha = a\,\sigma^\mu_{\alpha\dot\alpha}\xi^{\dagger\dot\alpha}\partial_\mu\varphi + b\,\xi_\alpha F \ , \quad \delta F = c\,\partial_\mu\psi^\alpha\sigma^\mu_{\alpha\dot\alpha}\xi^{\dagger\dot\alpha} \ .$$

The goal is now to match the couplings. We have:

$$\delta\mathcal{L}_\varphi = \partial_\mu\varphi^\dagger\partial^\mu(\delta\varphi) - \mu^2\varphi\,\delta\varphi + h.c.$$
$$= \partial_\mu\varphi^\dagger\xi^\alpha\partial^\mu\psi_\alpha - \mu^2\varphi\,\xi^\alpha\psi_\alpha + h.c.$$

$$\delta\mathcal{L}_\psi = \frac{1}{2}i\psi^\dagger\bar\sigma^\mu\partial_\mu(\delta\psi) - m\,\psi\delta\psi + h.c.$$
$$= \frac{1}{2}i\psi^\dagger_{\dot\beta}\bar\sigma^{\mu\dot\beta\alpha}\partial_\mu(a\,\sigma^\nu_{\alpha\dot\alpha}\xi^{\dagger\dot\alpha}\partial_\nu\varphi + b\,\xi_\alpha F) - m\,\psi^\alpha(a\,\sigma^\mu_{\alpha\dot\alpha}\xi^{\dagger\dot\alpha}\partial_\mu\varphi + b\,\xi_\alpha F) + h.c.$$
$$= \frac{1}{2}ia\,\psi^\dagger_{\dot\beta}\bar\sigma^{\mu\dot\beta\alpha}\sigma^\nu_{\alpha\dot\alpha}\xi^{\dagger\dot\alpha}\partial_\mu\partial_\nu\varphi + \frac{1}{2}ib\,\psi^\dagger\bar\sigma^\mu\xi\partial_\mu F - ma\,\psi\sigma^\mu\xi^\dagger\partial_\mu\varphi - mb\,\psi\xi F + h.c.$$
$$= \frac{1}{2}ia\,\psi^\dagger_{\dot\beta}(2\eta^{\mu\nu}\delta^{\dot\beta}_{\dot\alpha})\xi^{\dagger\dot\alpha}\partial_\mu\partial_\nu\varphi + 0 - ma\,\psi\sigma^\mu\xi^\dagger\partial_\mu\varphi - mb\,\psi\xi\,F + h.c.$$
$$= ia\,\psi^\dagger\xi^\dagger\partial^2\varphi - ma\,\psi\sigma^\mu\xi^\dagger\partial_\mu\varphi - mb\,\psi\xi\,F + h.c.$$

The 0 came from $\partial_\mu F = 0$, since F has no kinetic term and hence no dynamics. Since $i\psi^\dagger\xi^\dagger\partial^2\varphi + h.c. = -i\partial^\mu\psi^\dagger\xi^\dagger\partial_\mu\varphi + h.c. = +i\xi\partial^\mu\psi\partial_\mu\varphi^\dagger + h.c.$ up to a total derivative, we can write $\delta\mathcal{L}_\psi$ as
$$\delta\mathcal{L}_\psi = +ia\,\xi\partial^\mu\psi\partial_\mu\varphi^\dagger - ma\,\psi\sigma^\mu\xi^\dagger\partial_\mu\varphi - mb\,\psi\xi\,F + h.c.$$

The first term in $\delta\mathcal{L}_\varphi$ cancels the first term in $\delta\mathcal{L}_\psi$ if $a = +i$. (When verifying this, recall that $\xi\psi \equiv \xi^\alpha\psi_\alpha = -\psi_\alpha\xi^\alpha = +\psi^\alpha\xi_\alpha \equiv \psi\xi$.)

The variation of the $\varphi\psi$ interaction term is

$$\delta\mathcal{L}_{\varphi\psi} = g\,\delta\varphi\,\psi\psi + 2g\,\varphi\,\psi\delta\psi + h.c.$$
$$= g(\xi\psi)\psi\psi + 2g\,\varphi\,\psi(i\sigma^\mu\xi^\dagger\partial_\mu\varphi + b\xi F) + h.c.$$
$$= 0 + 2ig\,\varphi\,(\psi\sigma^\mu\xi^\dagger)\partial_\mu\varphi + 2gb\,\varphi\,(\psi\xi)F + h.c.$$

where the 0 comes from $\psi_\alpha(\psi\psi) = 0$. The variation of $\mathcal{L}_F$ is

$$\delta\mathcal{L}_F = F^\dagger\delta F - \kappa\,\delta F\varphi - \kappa\,F\delta\varphi - \lambda\delta F\varphi^2 - 2\lambda F\varphi\,\delta\varphi + h.c.$$
$$= (F^\dagger - \kappa\,\varphi - \lambda\,\varphi^2)\delta F - (\kappa + 2\lambda\,\varphi)F\,\delta\varphi + h.c.$$
$$= (F^\dagger - \kappa\,\varphi - \lambda\,\varphi^2)(c\,\partial_\mu\psi^\alpha\sigma^\mu_{\alpha\dot\alpha}\xi^{\dagger\dot\alpha}) - (\kappa + 2\lambda\,\varphi)F(\xi\psi) + h.c.$$
$$= 0 - (\kappa\,\varphi + \lambda\,\varphi^2)c\,\partial_\mu\psi\sigma^\mu\xi^\dagger - (\kappa + 2\lambda\,\varphi)F(\xi\psi) + h.c.$$

where the 0 comes from $F^\dagger\partial_\mu\psi\sigma^\mu\xi^\dagger = -\partial_\mu F^\dagger\,\psi\sigma^\mu\xi^\dagger + (\text{total derivative}) = 0$ since again $\partial_\mu F^\dagger = 0$.

Comparing the terms in $\delta\mathcal{L}_{\varphi\psi}$ and $\delta\mathcal{L}_F$, we want $gb = \lambda$ and $ig = -\lambda c$, where the minus sign arises from $\varphi^2\partial_\mu\psi\sigma^\mu\xi^\dagger = -2\varphi\,\partial_\mu\varphi\,\psi\sigma^\mu\xi^\dagger$ up to a total derivative. So we have

$$b = \frac{\lambda}{g} \quad \text{and} \quad c = -i\frac{g}{\lambda}\,.$$

At this point we should collect what we know so far. We have

$$\delta\mathcal{L} = -\mu^2\varphi\,\xi\psi - i\left(\frac{g}{\lambda}\kappa + m\right)\psi\sigma^\mu\xi^\dagger\partial_\mu\varphi - \left(\kappa + m\frac{\lambda}{g}\right)\psi\xi\,F$$

We therefore require $\mu^2 = 0$ and $\lambda = -\kappa g/m$. After rescaling $F \to mF/\kappa$ and then setting $\kappa = m$, we obtain

$$\mathcal{L} = \partial_\mu\varphi^\dagger\partial^\mu\varphi + i\psi^\dagger\bar{\sigma}^\mu\partial_\mu\psi + F^\dagger F + \left(-\frac{1}{2}m\psi\psi + g\varphi\psi\psi - (m - g\varphi)\varphi F + h.c.\right)$$

Upon the rephasing $\psi \to i\psi$ and letting $g \to -g$, we obtain the Lagrangian of equation (12) on p. 467.

2. In general there may be $\mathcal{N}$ supercharges Q_α^I, with $I = 1, ..., \mathcal{N}$. Show that we can have $\{Q_\alpha^I, Q_\beta^J\} = \varepsilon_{\alpha\beta}Z^{IJ}$, where Z^{IJ} denotes c-numbers known as central charges.

Solution:

Generalizing from p. 463, we have $\{Q_\alpha^I, Q_\beta^J\} = Z^{IJ}\varepsilon_{\alpha\beta}$ for some ordinary commuting numbers Z^{IJ}. For the case $\mathcal{N} = 1$, the fact that the left-hand side is symmetric in $\alpha\beta$ while the right-hand side is antisymmetric in $\alpha\beta$ forces $Z = 0$. If $\mathcal{N} > 1$ then we have more indices to play with; if the left-hand side is symmetric under the exchange $(\alpha, I) \leftrightarrow (\beta, J)$, then the right-hand side can also be symmetric and nonzero if $Z^{IJ} = -Z^{JI}$. So Z is an antisymmetric matrix whose components are ordinary c-numbers.

3. From the fact that we do not know how to write consistent quantum field theories with fields having spin greater than 2 show that the $\mathcal{N}$ in the previous exercise cannot exceed 8. Theories with $\mathcal{N} = 8$ supersymmetry are said to be maximally supersymmetric. Show that if we do not want to include gravity, $\mathcal{N}$ cannot be greater than 4. Supersymmetric $\mathcal{N} = 4$ Yang-Mills theory has many remarkable properties.

Solution:

To set the notation, consider $\mathcal{N} = 1$ supersymmetry. Let $|s\rangle$ denote a state of spin-s. The generator Q raises the spin by $+1/2$, so that $Q|s\rangle = |s + \frac{1}{2}\rangle$. Moreover, since Q is a

297

Grassmann generator we have $Q^2 = 0$. Thus, for example, starting with the state $s = 0$ we find the state $Q|0\rangle = |\frac{1}{2}\rangle$, and similarly $Q^\dagger|0\rangle = |-\frac{1}{2}\rangle$. Thus we find the chiral supermultiplet containing the states $s = 0$ and $s = \pm 1/2$.

Now generalize to the case $\mathcal{N} > 1$. The supersymmetry generators Q^I obtain an index $I = 1, 2, ..., \mathcal{N}$. Each Q^I increases the spin by $1/2$. Starting with a state $|s\rangle$, the highest spin we can obtain using the Q^I is given by the action of all of the Q^I together: $Q^1 Q^2 ... Q^\mathcal{N}|s\rangle = |s + \frac{1}{2}\mathcal{N}\rangle$.

If we begin from $s = -1$ and we want to restrict to spin-0, spin-1/2 and spin-1 states only, then we require $-1 + \frac{1}{2}\mathcal{N} \leq 1 \implies \mathcal{N} \leq 4$.

If we begin from $s = -2$ and demand that no states higher than spin-2 exist, then we require $-2 + \frac{1}{2}\mathcal{N} \leq 2 \implies \mathcal{N} \leq 8$.

4. Show that $\partial\theta_\alpha/\partial\theta^\beta = \varepsilon_{\alpha\beta}$.

Solution:

$$\frac{\partial}{\partial\theta^\beta}\theta_\alpha = \varepsilon_{\alpha\gamma}\frac{\partial}{\partial\theta^\beta}\theta^\gamma = \varepsilon_{\alpha\gamma}\delta^\gamma_\beta = \varepsilon_{\alpha\beta} \ .$$

Alternatively,

$$\frac{\partial}{\partial\theta^\beta}\theta_\alpha = \varepsilon_{\alpha\gamma}\frac{\partial}{\partial\theta^\beta}\theta^\gamma = \varepsilon_{\alpha\gamma}\int d\theta_\beta\theta^\gamma = \varepsilon_{\alpha\gamma}\delta^\gamma_\beta = \varepsilon_{\alpha\beta} \ .$$

5. Work out $\delta\varphi$, $\delta\psi$, and δF precisely by computing $\delta\Phi = i(\xi^\alpha Q_\alpha + \bar{\xi}_{\dot\alpha}\bar{Q}^{\dot\alpha})\Phi$.

Solution:

The chiral superfield is

$$\Phi = \varphi + \sqrt{2}\,\theta\psi + (\theta\theta)F + i(\theta\sigma^\mu\bar{\theta})\partial_\mu\varphi - \frac{1}{2}(\theta\sigma^\mu\bar{\theta})(\theta\sigma^\nu\bar{\theta})\partial_\mu\partial_\nu\varphi + i\sqrt{2}\,\theta(\theta\sigma^\mu\bar{\theta})\partial_\mu\psi$$

where as explained in the text, the field on the left depends on x and $\bar{\theta}$ only in the combination $y \equiv x + i\theta\sigma\bar{\theta}$, and it depends separately on θ, while the fields on the right depend only on x. The variation is therefore

$$\delta\Phi = \delta\varphi + \sqrt{2}\,\theta\delta\varphi + (\theta\theta)\delta F + ...$$

So when computing the action of the supercharges on Φ, we need to extract the term without any θ or $\bar{\theta}$ to get $\delta\varphi$, the term of order θ to get $\delta\psi$, and the term of order $(\theta\theta)$ to get δF.

The supercharges are

$$Q_\alpha = +\frac{\partial}{\partial\theta^\alpha} - i\sigma^\mu_{\alpha\dot\alpha}\bar\theta^{\dot\alpha}\partial_\mu$$

$$\bar Q_{\dot\alpha} = -\frac{\partial}{\partial\bar\theta^{\dot\alpha}} + i\theta^\alpha\sigma^\mu_{\alpha\dot\alpha}\partial_\mu$$

Direct computation yields

$$Q_\alpha\Phi = \sqrt{2}\,\psi_\alpha + i\sqrt{2}\,(\theta\sigma^\mu\bar\theta)\partial_\mu\psi_\alpha + \sqrt{2}\,(\bar\theta\bar\theta)\theta_\alpha(\theta\partial^2\psi) - i\sigma^\mu_{\alpha\dot\alpha}\bar\theta^{\dot\alpha}(\theta\theta)\partial_\mu F - 2\theta_\alpha F$$

$$\bar Q^{\dot\alpha}\Phi = -\bar\sigma^{\mu\dot\alpha\alpha}\theta_\alpha\partial_\mu\varphi - 2\sqrt{2}\,i\,\bar\sigma^{\mu\dot\alpha\alpha}\theta_\alpha(\theta\partial_\mu\psi)$$

so that the variation $\delta\Phi = i(\xi^\alpha Q_\alpha + \bar\xi_{\dot\alpha}\bar Q^{\dot\alpha})\Phi$ gives

$$\delta\Phi = i\sqrt{2}\,\xi^\alpha\psi_\alpha + \bar\xi_{\dot\alpha}\bar\sigma^{\mu\dot\alpha\alpha}\theta_\alpha\partial_\mu\varphi - 2i\theta_\alpha F + 2\sqrt{2}\,\bar\xi_{\dot\alpha}\bar\sigma^{\mu\dot\alpha\alpha}\theta_\alpha\,\theta^\beta\partial_\mu\psi_\beta + ...$$

Since $\theta_\alpha\theta^\beta = \frac{1}{2}(\theta\theta)\delta_\alpha^\beta$, we have

$$\delta\Phi = i\sqrt{2}\,\xi^\alpha\psi_\alpha + \bar\xi_{\dot\alpha}\bar\sigma^{\mu\dot\alpha\alpha}\theta_\alpha\partial_\mu\varphi - 2i\theta_\alpha F + \sqrt{2}\,\bar\xi_{\dot\alpha}\bar\sigma^{\mu\dot\alpha\alpha}(\theta\theta)\partial_\mu\psi_\alpha + ...$$

Matching the terms order by order in θ as described above gives

$$\delta\varphi = i\sqrt{2}\,\xi^\alpha\psi_\alpha\ ,\ \ \delta\psi^\alpha = -\frac{1}{\sqrt{2}}\bar\xi_{\dot\alpha}\bar\sigma^{\mu\dot\alpha\alpha}\partial_\mu\varphi + i\sqrt{2}\,\xi^\alpha F\ ,\ \ \delta F = \sqrt{2}\,\bar\xi_{\dot\alpha}\bar\sigma^{\mu\dot\alpha\alpha}\partial_\mu\psi_\alpha\ .$$

If instead you desire $\delta\psi_\alpha = \varepsilon_{\alpha\beta}\delta\psi^\beta$, then use

$$\varepsilon_{\alpha\beta}\bar\xi_{\dot\alpha}\bar\sigma^{\mu\dot\alpha\beta} = \varepsilon_{\alpha\beta}\varepsilon_{\dot\alpha\dot\beta}\bar\xi^{\dot\beta}\bar\sigma^{\mu\dot\alpha\beta} = -\varepsilon_{\alpha\beta}\varepsilon_{\dot\beta\dot\alpha}\bar\xi^{\dot\beta}\bar\sigma^{\mu\dot\alpha\beta} = -\bar\xi^{\dot\beta}\sigma^\mu_{\alpha\dot\beta}$$

to get $\delta\psi_\alpha = +\frac{1}{\sqrt{2}}\partial_\mu\varphi\sigma^\mu_{\alpha\dot\alpha}\bar\xi^{\dot\alpha} + i\sqrt{2}\,\xi_\alpha F$, which is the form given under equation (9) on p. 465.

6. For any polynomial $W(\Phi)$ show that $[W(\Phi)]_F = F[dW(\varphi)/d\varphi]+$ terms not involving F. Show that for the theory (11) the potential energy is given by $V(\varphi^\dagger,\varphi) = |\partial W(\varphi)/\partial\varphi|^2$.

Solution:

This will follow from the binomial expansion

$$(p+q)^n = \sum_{j=0}^{n}\frac{n!}{j!(n-j)!}\,p^j q^{n-j}\ .$$

A chiral superfield has the expansion $\Phi = \varphi + \sqrt{2}\,\theta\psi + (\theta\theta)F + ...$, where the (...) are terms that will not contribute to the F term of $W(\Phi)$. From now on we drop the (...). Any polynomial $W(\Phi)$ of degree D can be written as $W(\Phi) = \sum_{n=0}^{D}C_n\Phi^n$, where C_n are constants whose

numerical values are unimportant for this problem. Let $p = (\theta\theta)F$ and $q = \varphi + \sqrt{2}\,\theta\psi$ in the binomial expansion:

$$\Phi^n = \sum_{j=0}^{n} \frac{n!}{j!(n-j)!}\,[(\theta\theta)F]^j(\varphi + \sqrt{2}\,\theta\psi)^{n-j}$$

Since $(\theta\theta)^2 = 0$, the series terminates at $j = 1$:

$$\Phi^n = (\varphi + \sqrt{2}\,\theta\psi)^n + n(\theta\theta)F(\varphi + \sqrt{2}\,\theta\psi)^{n-1}$$

Moreover, since $(\theta\theta)\theta^{n-1} = 0$ for $n > 1$, we have

$$\Phi^{n>1} = (\varphi + \sqrt{2}\,\theta\psi)^n + n(\theta\theta)F\varphi^{n-1}$$

We could continue to simplify the first term, but we are asked to consider only the terms that involve F. So as far as this problem is concerned, we are free to write

$$(\Phi^{n>1})_F = nF\varphi^{n-1} + \ldots$$

Since $n\varphi^{n-1} = \frac{d}{d\varphi}\varphi^n$, we are just about done:

$$W(\Phi) = \sum_{n=0}^{D} C_n\Phi^n = \sum_{n=0}^{D} C_n nF\varphi^{n-1} + \ldots = F\frac{d}{d\varphi}\sum_{n=0}^{D} C_n\varphi^n = F\frac{d}{d\varphi}W(\varphi) + \ldots$$

where the (...) does not involve the field F. Now integrate out the auxiliary fields $\{F_a\}_{a=1}^3$ using the Gaussian identity

$$\int \mathcal{D}F\mathcal{D}F^\dagger\, e^{i\int d^4x(F^\dagger F + J^\dagger F + F^\dagger J)} = e^{-i\int d^4x\, J^\dagger(x)J(x)}$$

where $J^\dagger = \frac{d}{d\varphi}W(\varphi)$ and $J = (J^\dagger)^\dagger = [\frac{d}{d\varphi}W(\varphi)]^*$. After integrating out the F field, the Lagrangian gets the term

$$\mathcal{L}_W = -J^\dagger J = -\left[\frac{dW(\varphi)}{d\varphi}\left(\frac{dW(\varphi)}{d\varphi}\right)^*\right] = -\left|\frac{dW(\varphi)}{d\varphi}\right|^2$$

The potential energy is $V = -\mathcal{L}_W = |dW(\varphi)/d\varphi|^2$.

7. Construct a field theory in which supersymmetry is spontaneously broken. [Hint: You need at least three chiral superfields.]

Solution:

Let $\Xi(x,\theta) = \xi(x) + \ldots$ and $\{\Phi_i(x,\theta) = \varphi_i(x) + \ldots\}_{i=1}^n$ be chiral superfields, and consider a superpotential of the form[46]

$$W(\Xi, \Phi) = \sum_{i=1}^n \Phi_i f_i(\Xi)$$

with the $f_i(\Xi)$ being as-of-yet unspecified functions of Ξ.

Supersymmetry is spontaneously broken at tree level if there exist values of the scalar components $\xi(x)$ and $\varphi_i(x)$ such that the potential is zero. From the previous problem, we know that this is equivalent to the conditions

$$\frac{\partial W(\xi,\varphi)}{\partial \xi} = \sum_{i=1}^n \varphi_i f_i'(\xi) = 0 \qquad \text{and} \qquad \frac{\partial W(\xi,\varphi)}{\partial \varphi_i} = f_i(\xi) = 0 \;.$$

The first condition is satisfied by the field values $\varphi_i = 0$. The second equation, however, imposes n conditions on a function of just one variable, ξ. If $n \geq 2$, then these conditions cannot be satisfied and supersymmetry is spontaneously broken.

For example, take $n = 2$ and consider the functions $f_1(\xi) = \xi - \lambda$ and $f_2(\xi) = \xi^2$, with λ an arbitrary nonzero constant. Then $\xi = \lambda$ sets $f_1(\xi) = 0$ but $f_2(\xi) \neq 0$, while $\xi = 0$ sets $f_2(\xi) = 0$ but $f_1(\xi) \neq 0$.

8. If we can construct supersymmetric quantum field theory, surely we can construct supersymmetric quantum mechanics. Indeed, consider $Q_1 \equiv \frac{1}{2}[\sigma_1 P + \sigma_2 W(x)]$ and $Q_2 \equiv \frac{1}{2}[\sigma_2 P - \sigma_1 W(x)]$, where the momentum operator $P = -i(d/dx)$ as usual. Define $Q \equiv Q_1 + iQ_2$. Study the properties of the Hamiltonian H defined by $\{Q, Q^\dagger\} = 2H$.

Solution:

We will briefly follow section 2 of arXiv:hep-th/9405029v2. Consult this reference for an extensive review.

As described in the problem, it is possible to factorize a Hamiltonian $H = -\frac{1}{2}\frac{d^2}{dx^2} + V(x)$ as $H = h^\dagger h$, where $h \equiv \frac{1}{\sqrt{2}}\frac{d}{dx} + W(x)$ and $h^\dagger \equiv -\frac{1}{\sqrt{2}}\frac{d}{dx} + W(x)$. In terms of these variables, and using $\frac{d}{dx}[W(x)f(x)] - W(x)\frac{df}{dx} = \frac{dW}{dx}f(x)$, we find $V(x) = W(x)^2 - \frac{1}{\sqrt{2}}W'(x)$. If the ground state has zero energy, then a state ψ_0 that satisfies $h\psi = 0$ automatically satisfies $H\psi = 0$ and thus is the ground state. Solving $A\psi_0 = 0$ gives $W(x) = -\frac{1}{\sqrt{2}}\frac{\psi_0'(x)}{\psi_0(x)}$. Thus if we know $\psi_0(x)$, then we know $W(x)$ and $V(x)$.

If we define a new Hamiltonian $\tilde{H} \equiv hh^\dagger$ and put it into the form $\tilde{H} = -\frac{1}{2}\frac{d^2}{dx^2} + \tilde{V}(x)$,

---

[46]Here we follow Section 26.5 in Weinberg, Volume III. The original paper is L. O'Raifeartaigh, "Spontaneous symmetry breaking for chiral scalar superfields," Nucl. Phys. B96 (1975) 331-352.

then by direct computation we find $\tilde{V}(x) = W(x)^2 + \frac{1}{\sqrt{2}}W'(x)$. The reason for defining this new Hamiltonian is the following.

Let ψ be an eigenstate of H with eigenvalue E, meaning $H\psi = E\psi$. Then $\tilde{H}(h\psi) = hh^\dagger h\psi = hH\psi = hE\psi = E(h\psi)$. Therefore, if ψ is an eigenstate of H with eigenvalue E, then $h\psi$ is an eigenstate of $\tilde{H}$ with eigenvalue E. Similarly, if $\tilde{\psi}$ is an eigenstate of $\tilde{H}$ with eigenvalue $\tilde{E}$, then $h^\dagger \tilde{\psi}$ is an eigenstate of H with eigenvalue $\tilde{E}$.

Therefore we can solve $H\psi = E\psi$ by solving $\tilde{H}\tilde{\psi} = \tilde{E}\tilde{\psi}$, or vice versa.

The underlying reason for this is that the two Hamiltonians belong to a supersymmetry algebra. The operators

$$\mathcal{H} \equiv \begin{pmatrix} H & 0 \\ 0 & \tilde{H} \end{pmatrix} , \quad Q \equiv \begin{pmatrix} 0 & 0 \\ h & 0 \end{pmatrix} , \quad Q^\dagger \equiv \begin{pmatrix} 0 & h^\dagger \\ 0 & 0 \end{pmatrix}$$

satisfy the algebra $[\mathcal{H}, Q] = [\mathcal{H}, Q^\dagger] = \{Q, Q\} = \{Q^\dagger, Q^\dagger\} = 0$ and $\{Q, Q^\dagger\} = \mathcal{H}$.

An important property of this Hamiltonian $\mathcal{H}$ is that the energy of any state is non-negative. A state $|\text{inv}\rangle$ invariant under supersymmetry satisfies $Q_i|\text{inv}\rangle = 0$. A state $|\text{not}\rangle$ that is not invariant under supersymmetry satisfies $Q_i|\text{not}\rangle \neq 0$. Since $\mathcal{H}$ is the sum of squares of Hermitian operators, its eigenvalues are non-negative, so $Q_i|\text{not}\rangle > 0$. Since all states besides $|\text{inv}\rangle$ have energy greater than zero, the state $|\text{inv}\rangle$ is the vacuum state. This tells us that if a supersymmetric state exists, it must be the vacuum state of the theory.

This has implications for trying to break supersymmetry spontaneously. For further details, consult Witten, "Dynamical Breaking of Supersymmetry," Nucl. Phys. B185 (1981) 513-554.

IX   Part N

IX.1   N.2 Gluon Scattering in Pure Yang-Mills Theory

1. Work out the two polarization vectors for general μ and $\tilde{\mu}$ for a gluon moving along the third direction.

Solution:

Let $p^\mu = E(1,0,0,1)$ be the momentum of a gluon moving along the third spatial direction. The 2-by-2 matrix representation of p^μ is

$$p_{\alpha\dot{\alpha}} = \sigma^\mu_{\alpha\dot{\alpha}} p_\mu = E(\sigma^0_{\alpha\dot{\alpha}} - \sigma^3_{\alpha\dot{\alpha}}) = 2E \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} .$$

Therefore we have

$$\lambda_\alpha = \sqrt{2E} \begin{pmatrix} 0 \\ 1 \end{pmatrix} , \quad \tilde{\lambda}_{\dot{\alpha}} = \sqrt{2E} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

as on p. 487. The polarization vectors (in 2-by-2 matrix notation) are given on p. 489:

$$\epsilon^+_{\alpha\dot\alpha} = \frac{\mu_\alpha \tilde\lambda_{\dot\alpha}}{\langle\mu\lambda\rangle} \ , \quad \epsilon^-_{\alpha\dot\alpha} = \frac{\lambda_\alpha \tilde\mu_{\dot\alpha}}{[\lambda\mu]}$$

where $\langle\mu\lambda\rangle \equiv \varepsilon^{\alpha\beta}\mu_\alpha\lambda_\beta$ and $[\lambda\mu] \equiv \varepsilon^{\dot\alpha\dot\beta}\tilde\lambda_{\dot\alpha}\tilde\mu_{\dot\beta}$. Using the explicit forms for λ and $\tilde\lambda$, we have

$$\langle\mu\lambda\rangle = \varepsilon^{12}\mu_1\lambda_2 + 0 = \sqrt{2E}\,\mu_1$$

and

$$[\lambda\mu] = 0 + \varepsilon^{\dot2\dot1}\tilde\lambda_{\dot2}\tilde\mu_{\dot1} = \sqrt{2E}\,\tilde\mu_{\dot1} \ .$$

The numerators are

$$\mu_\alpha\tilde\lambda_{\dot\alpha} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}(0, \sqrt{2E}) = \sqrt{2E}\begin{pmatrix} 0 & \mu_1 \\ 0 & \mu_2 \end{pmatrix}$$

and

$$\lambda_\alpha\tilde\mu_{\dot\alpha} = \begin{pmatrix} 0 \\ \sqrt{2E} \end{pmatrix}(\tilde\mu_{\dot1}, \tilde\mu_{\dot2}) = \sqrt{2E}\begin{pmatrix} 0 & 0 \\ \tilde\mu_{\dot1} & \tilde\mu_{\dot2} \end{pmatrix} \ .$$

Therefore

$$\epsilon^+_{\alpha\dot\alpha} = \begin{pmatrix} 0 & 1 \\ 0 & \frac{\mu_2}{\mu_1} \end{pmatrix} \ , \quad \epsilon^-_{\alpha\dot\alpha} = \begin{pmatrix} 0 & 0 \\ 1 & \frac{\tilde\mu_{\dot2}}{\tilde\mu_{\dot1}} \end{pmatrix} \ .$$

3. Show that the result in (17) satisfies the reflection identity $A(1^-, 2^-, 3^+, 4^+) = A(4^+, 3^+, 2^-, 1^-)$.

$$A(1^-, 2^-, 3^+, 4^+) = \frac{\langle 12\rangle^4}{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle} = \frac{p_1 \cdot p_2}{p_2 \cdot p_3} \qquad (17)$$

Solution:

The Parke-Taylor formula in equation (20) on p. 493 implies

$$A(a^+, b^+, c^-, d^-) = \frac{\langle cd\rangle^4}{\langle ab\rangle\langle bc\rangle\langle cd\rangle\langle da\rangle} \ .$$

Let $a = 4, b = 3, c = 2, d = 1$ to get

$$A(4^+, 3^+, 2^-, 1^-) = \frac{\langle 21\rangle^4}{\langle 43\rangle\langle 32\rangle\langle 21\rangle\langle 14\rangle} = \frac{(-1)^4\langle 12\rangle^4}{(-1)^4\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle} = A(1^-, 2^-, 3^+, 4^+) \ \checkmark$$

4. Show that the "all plus" and "all minus" cubic Yang-Mills vertices (see appendix 2) vanish. [Hint: Choose the μ spinors wisely.]

Solution:

The color-stripped all-plus cubic amplitude is (see equation (23) on p. 495)

$$A(1^+, 2^+, 3^+) = (\epsilon_1^+ \cdot \epsilon_2^+)(\epsilon_3^+ \cdot p_1) + (\epsilon_2^+ \cdot \epsilon_3^+)(\epsilon_1^+ \cdot p_2) + (\epsilon_3^+ \cdot \epsilon_1^+)(\epsilon_2^+ \cdot p_3) \,.$$

Referencing p. 490, we have

$$\epsilon_i^+ \cdot \epsilon_j^+ = \frac{\langle \mu_i \mu_j \rangle [\lambda_i \lambda_j]}{\langle \mu_i \lambda_i \rangle \langle \mu_j \lambda_j \rangle} \quad \text{and} \quad \epsilon_i^+ \cdot p_j = \frac{\langle \mu_i \lambda_j \rangle [\lambda_i \lambda_j]}{\langle \mu_i \lambda_i \rangle} \,.$$

If $\mu_1 = \mu_2 = \mu_3$, then all $\epsilon_i^+ \cdot \epsilon_j^+ = 0$. Therefore $A(1^+, 2^+, 3^+) = 0$. Flipping all three helicities from $+$ to $-$ will make each term in the amplitude have a factor of $\epsilon_i^- \cdot \epsilon_j^- \propto [\mu_i \mu_j]$. If we choose all $\tilde{\mu}_i$ equal, then $A(1^-, 2^-, 3^-) = 0$.

5. Why doesn't the argument in the text that $A(-++...++)$ vanish apply to $A(-++)$?

Solution:

For the 3-point amplitude, either all of the $\langle ij \rangle = 0$ or all of the $[ij] = 0$ (see equation (17) on p. 508). Therefore there are zeros in the denominator of the amplitude, so that choosing the μ spinors wisely yields an indeterminate zero over zero situation rather than just zero. Indeed, in appendix 2 on p. 496 we see the μ factors cancel out explicitly, leaving behind the vertex (25) or (26).

6. Insert the expansion for the cubic vertex into (21) and derive $M(W_1^+, W_2^+, Z_3^-)$.

$$M(W_i) = \int d^2\lambda_i \, e^{i\tilde{\mu}_i^\alpha \lambda_{i\alpha}} M(\lambda_i, \tilde{\lambda}_i) \qquad\qquad (21a)$$

$$M(Z_i) = \int d^2\tilde{\lambda}_i \, e^{i\mu_i^{\dot\alpha} \tilde{\lambda}_{i\dot\alpha}} M(\lambda_i, \tilde{\lambda}_i) \qquad\qquad (21b)$$

Solution:

We follow a lecture given by A. Zee and documented by Oscar Castillo-Felisola.

Recall that the amplitudes M are related to the amplitudes A by including the momentum-conserving delta function:

$$M(1, ..., n) = A(1, ..., n) \, \delta^{(4)}\left(\sum_{i=1}^n \lambda_i \tilde{\lambda}_i \right) \,.$$

The cubic vertex is

$$A(1^+, 2^+, 3^-) = \frac{[12]^3}{[23][31]} \,.$$

We can write a fourier representation of the delta function

$$\delta^{(4)}\left(\sum_{i=1}^{n}\lambda^{(i)}\tilde{\lambda}^{(i)}\right) = \frac{1}{(2\pi)^4}\int d^4x\, e^{ix^{\alpha\dot\alpha}\sum_i \lambda_{i\alpha}\tilde\lambda_{i\dot\alpha}}$$

and perform the relevant twistor transforms, using W for $+$ and Z for $-$. The $(2\pi)^4$ will cancel from the resulting fourier transforms, so from now on we will simply drop the factors of 2π. We have:

$$M(W_1^+, W_2^+, Z_3^-) = \int d^2\lambda_1 d^2\lambda_2 d^2\tilde\lambda_3\, e^{i(\tilde\mu_1\lambda_1+\tilde\mu_2\lambda_2+\mu_3\tilde\lambda_3)}\frac{[12]^3}{[23][31]}\int d^4x\, e^{ix^{\alpha\dot\alpha}\sum_i \lambda_{i\alpha}\tilde\lambda_{i\dot\alpha}}$$

$$= [12]^3\int d^4x\left(\int d^2\lambda_1\, e^{i(\tilde\mu_1+x\tilde\lambda_1)\lambda_1}\right)\left(\int d^2\lambda_2\, e^{i(\tilde\mu_2+x\tilde\lambda_2)\lambda_2}\right)\int d^2\tilde\lambda_3\frac{e^{i(\mu_3+x\lambda_3)\tilde\lambda_3}}{[23][31]}$$

$$= [12]^3\int d^4x\left(\delta^2(\tilde\mu_1+x\tilde\lambda_1)\right)\left(\delta^2(\tilde\mu_2+x\tilde\lambda_2)\right)\int d^2\tilde\lambda_3\frac{e^{i(\mu_3+x\lambda_3)\tilde\lambda_3}}{[23][31]}\ .$$

Since two spinors span the 2-dimensional spinor space, we can write a third spinor as a linear combination of the first two: $\tilde\lambda_3 = a_1\tilde\lambda_1+a_2\tilde\lambda_2$. The goal is to perform a change of integration variables from $d^2\tilde\lambda_3$ to $da_1 da_2$.

The reader seeing this calculation for the first time may be distracted by the indices and tildes, but in fact the change of variables is completely straightforward. Let

$$f_{\dot\alpha} \equiv \frac{\partial(\tilde\lambda_3)_{\dot\alpha}}{\partial a_1} = (\tilde\lambda_1)_{\dot\alpha}\ ,\quad g_{\dot\alpha} \equiv \frac{\partial(\tilde\lambda_3)_{\dot\alpha}}{\partial a_2} = (\tilde\lambda_2)_{\dot\alpha}\ .$$

Then we have

$$d(\tilde\lambda_3)_{\dot\alpha} = f_{\dot\alpha}da_1 + g_{\dot\alpha}da_2$$

and therefore, just as in the usual calculus,

$$d^2\tilde\lambda_3 = |f_1 g_2 - g_1 f_2|da_1 da_2$$
$$= |\varepsilon^{\dot\alpha\dot\beta}(\tilde\lambda_1)_{\dot\alpha}(\tilde\lambda_2)_{\dot\beta}|da_1 da_2$$
$$= |[12]|\, da_1 da_2\ .$$

Thus we find $d^2\tilde\lambda_3 = |[12]|da_1 da_2 = [12]\text{sgn}([12])da_1 da_2$. We also have $[13] = a_2[12]$ and $[32] = a_1[12]$ from writing the linear dependence equation as $|3] = a_1|1]+a_2|2]$ and contracting with the appropriate spinors. Therefore, the amplitude is

$$M(W_1^+, W_2^+, Z_3^-) = [12]^2\text{sgn}([12])\int d^4x\left(\int\frac{da_1}{a_1}e^{i(\mu_3+x\lambda_3)\tilde\lambda_1 a_1}\right)\left(\int\frac{da_2}{a_2}e^{i(\mu_3+x\lambda_3)\tilde\lambda_2 a_2}\right)$$
$$\times \delta^2(\tilde\mu_1+x\tilde\lambda_1)\delta^2(\tilde\mu_2+x\tilde\lambda_2)$$

Using the property $\delta(ax) = \frac{1}{|a|}\delta(x)$, the delta functions bring two powers of $[12]$ in the denominator, which cancels the factor of $[12]^2$ in the numerator. In the exponentials, the

305

delta functions set $x\tilde{\lambda}_1 = -\tilde{\mu}_1$ and $x\tilde{\lambda}_2 = -\tilde{\mu}_2$. We also recognize the integral representation of the sign function as $\text{sgn}(x) = \int \frac{dy}{y} e^{ixy}$. We have:

$$M(W_1^+, W_2^+, Z_3^-) = \text{sgn}([12])\,\text{sgn}(\mu_3\tilde{\lambda}_1 - \tilde{\mu}_1\lambda_3)\,\text{sgn}(\mu_3\tilde{\lambda}_2 - \tilde{\mu}_2\lambda_3)$$

We recognize the Lorentz invariant products as $W_1 I W_2$, $W_1 \cdot Z_3$ and $W_2 \cdot Z_3$. Thus we obtain the result

$$M(W_1^+, W_2^+, Z_3^-) = \text{sgn}(W_1 I W_2)\,\text{sgn}(W_1 \cdot Z_3)\,\text{sgn}(W_2 \cdot Z_3)\;.$$

7. Show that $M(W_1^+, Z_2^-, W_3^+, Z_4^-)$ reproduces (19).

$$A(1^-, 2^+, 3^-, 4^+) = \frac{\langle 13 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle} \tag{19}$$

Solution:

To clarify, we hope to reproduce (19) but with the opposite helicity:

$$A(1^+, 2^-, 3^+, 4^-) = \frac{[13]^4}{[12][23][34][41]}$$

The amplitude is given on p. 495 as

$$M(W_1^+, Z_2^-, W_3^+, Z_4^-) = \text{sgn}(W_1 \cdot Z_2)\text{sgn}(Z_2 \cdot W_3)\text{sgn}(W_3 \cdot Z_4)\text{sgn}(Z_4 \cdot W_1)$$

and the inverse transforms for $W = (\tilde{\mu}, \tilde{\lambda})$ and $Z = (\lambda, \mu)$ are

$$M(W_i) = \int d^2\lambda_i\, e^{i\tilde{\mu}_i^\alpha \lambda_{i\alpha}} M(\lambda_i, \tilde{\lambda}_i) \implies M(\lambda_i, \tilde{\lambda}_i) = \int \frac{d^2\tilde{\mu}_i}{(2\pi)^2} e^{-i\tilde{\mu}_i^\alpha \lambda_{i\alpha}} M(W_i)$$

$$M(Z_i) = \int d^2\tilde{\lambda}_i\, e^{i\mu_i^{\dot\alpha} \tilde{\lambda}_{i\dot\alpha}} M(\lambda_i, \tilde{\lambda}_i) \implies M(\lambda_i, \tilde{\lambda}_i) = \int \frac{d^2\mu_i}{(2\pi)^2} e^{-i\mu_i^{\dot\alpha} \tilde{\lambda}_{i\dot\alpha}} M(Z_i)$$

The integral representation of the sign function is $\text{sgn}(x) = \int \frac{da}{a} e^{iax}$. The amplitude is

$$M(W_1^+, Z_2^-, W_3^+, Z_4^-) = \int \frac{da_1}{a_1} \frac{da_2}{a_2} \frac{da_3}{a_3} \frac{da_4}{a_4}\, e^{i(a_1 W_1 \cdot Z_2 + a_2 Z_2 \cdot W_3 + a_3 W_3 \cdot Z_4 + a_4 Z_4 \cdot W_1)}\;.$$

306

Inverse twistor transforming gives

$$M(1^+, 2^-, 3^+, 4^-) = \int \frac{d^2\tilde{\mu}_1}{(2\pi)^2} \, e^{-i\tilde{\mu}_1\lambda_1} \int \frac{d^2\mu_2}{(2\pi)^2} \, e^{-i\mu_2\tilde{\lambda}_2} \int \frac{d^2\tilde{\mu}_3}{(2\pi)^2} \, e^{-\tilde{\mu}_3\lambda_3} \int \frac{d^2\mu_4}{(2\pi)^2} \, e^{-i\mu_4\tilde{\lambda}_4}$$

$$\times \int \frac{da_1}{a_1} \frac{da_2}{a_2} \frac{da_3}{a_3} \frac{da_4}{a_4} \, e^{i[a_1(\tilde{\mu}_1\lambda_2 - \tilde{\lambda}_1\mu_2) + a_2(\lambda_2\tilde{\mu}_3 - \mu_2\tilde{\lambda}_3) + a_3(\tilde{\mu}_3\lambda_4 - \tilde{\lambda}_3\mu_4) + a_4(\lambda_4\tilde{\mu}_1 - \mu_4\tilde{\lambda}_1)]}$$

$$= \int \frac{da_1}{a_1} \cdots \frac{da_4}{a_4} \int \frac{d^2\tilde{\mu}_1}{(2\pi)^2} \, e^{i(-\lambda_1 + a_1\lambda_2 + a_4\lambda_4)\tilde{\mu}_1} \int \frac{d^2\mu_2}{(2\pi)^2} \, e^{-i(\tilde{\lambda}_2 + a_1\tilde{\lambda}_1 + a_2\tilde{\lambda}_3)\mu_2}$$

$$\times \int \frac{d^2\tilde{\mu}_3}{(2\pi)^2} \, e^{i(-\lambda_3 + a_2\lambda_2 + a_3\lambda_4)\tilde{\mu}_3} \int \frac{d^2\mu_4}{(2\pi)^2} \, e^{-i(\tilde{\lambda}_4 + a_3\tilde{\lambda}_3 + a_4\tilde{\lambda}_1)\mu_4}$$

$$= \int \frac{da_1}{a_1} \cdots \frac{da_4}{a_4} \, \delta^2(-\lambda_1 + a_1\lambda_2 + a_4\lambda_4) \, \delta^2(\tilde{\lambda}_2 + a_1\tilde{\lambda}_1 + a_2\tilde{\lambda}_3)$$

$$\times \, \delta^2(-\lambda_3 + a_2\lambda_2 + a_3\lambda_4) \, \delta^2(\tilde{\lambda}_4 + a_3\tilde{\lambda}_3 + a_4\tilde{\lambda}_1) \ .$$

Now we need the "inverse" of the steps in the previous problem, namely to turn an integral $\int da_i/a_i \int da_j/a_j$ into $\int d^2\lambda$.

The first delta function sets $\lambda_1 = a_1\lambda_2 + a_4\lambda_4$, or in other words $|1\rangle = a_1|2\rangle + a_4|4\rangle$. Multiplying on the left with $\langle 2|$ gives $\langle 21 \rangle = a_4\langle 24 \rangle \implies a_4 = \frac{\langle 21 \rangle}{\langle 24 \rangle}$. Instead multiplying on the left with $\langle 4|$ gives $\langle 41 \rangle = a_1\langle 42 \rangle \implies a_1 = \frac{\langle 41 \rangle}{\langle 42 \rangle}$. As in N.2.6, we have $d^2\lambda_1 = |\langle 24 \rangle| \, da_1 da_4$ and therefore

$$\frac{da_1}{a_1} \frac{da_4}{a_4} = \frac{|\langle 24 \rangle|}{\langle 12 \rangle \langle 41 \rangle} \, d^2\lambda_1 \ .$$

We can therefore perform the integrals over a_1 and a_4 to get

$$M(1^+, 2^-, 3^+, 4^-) = \int \frac{da_2}{a_2} \frac{da_3}{a_3} \frac{|\langle 24 \rangle|}{\langle 12 \rangle \langle 41 \rangle} \, \delta^2\left(\tilde{\lambda}_2 + \frac{\langle 41 \rangle}{\langle 42 \rangle} \tilde{\lambda}_1 + a_2\tilde{\lambda}_3\right)$$

$$\times \, \delta^2(-\lambda_3 + a_2\lambda_2 + a_3\lambda_4) \, \delta^2\left(\tilde{\lambda}_4 + a_3\tilde{\lambda}_3 + \frac{\langle 21 \rangle}{\langle 24 \rangle} \tilde{\lambda}_1\right) \ .$$

Now repeat the previous procedure to perform the integral over a_2 and a_3. The middle delta function in the above expression sets $\lambda_3 = a_2\lambda_2 + a_3\lambda_4$, or in other words $|3\rangle = a_2|2\rangle + a_3|4\rangle$. Multiplying on the left with $\langle 2|$ implies $\langle 23 \rangle = a_3\langle 24 \rangle \implies a_3 = \frac{\langle 23 \rangle}{\langle 24 \rangle}$. Multiplying with $\langle 4|$ implies $\langle 43 \rangle = a_2\langle 42 \rangle \implies a_2 = \frac{\langle 43 \rangle}{\langle 42 \rangle}$. Using $d^2\lambda_3 = |\langle 24 \rangle| \, da_2 \, da_3$, we get

$$\frac{da_2}{a_2} \frac{da_3}{a_3} = \frac{|\langle 24 \rangle|}{\langle 23 \rangle \langle 34 \rangle} \, d^2\lambda_3 \ .$$

Therefore, we have

$$M(1^+, 2^-, 3^+, 4^-) = \left(\frac{|\langle 24 \rangle|}{\langle 12 \rangle \langle 41 \rangle}\right)\left(\frac{|\langle 24 \rangle|}{\langle 23 \rangle \langle 34 \rangle}\right) \delta^2\left(\tilde{\lambda}_2 + \frac{\langle 41 \rangle}{\langle 42 \rangle}\tilde{\lambda}_1 + \frac{\langle 43 \rangle}{\langle 42 \rangle}\tilde{\lambda}_3\right) \delta^2\left(\tilde{\lambda}_4 + \frac{\langle 23 \rangle}{\langle 24 \rangle}\tilde{\lambda}_3 + \frac{\langle 21 \rangle}{\langle 24 \rangle}\tilde{\lambda}_1\right)$$

$$= \frac{\langle 24 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle} \, \delta^2(\langle 24 \rangle\tilde{\lambda}_2 - \langle 41 \rangle\tilde{\lambda}_1 - \langle 43 \rangle\tilde{\lambda}_3) \, \delta^2(\langle 24 \rangle\tilde{\lambda}_4 + \langle 23 \rangle\tilde{\lambda}_3 + \langle 21 \rangle\tilde{\lambda}_1)$$

The hope is that the two delta functions recollect into $\delta^4(\sum_{i=1}^4 p_i)$. The equation $\sum_{i=1}^4 p_i = 0$ is written in terms of spinors as $\lambda_1 \tilde{\lambda}_1 + \lambda_2 \tilde{\lambda}_2 + \lambda_3 \tilde{\lambda}_3 + \lambda_4 \tilde{\lambda}_4 = 0$. Contracting with λ_4 implies $\langle 41 \rangle \tilde{\lambda}_1 + \langle 42 \rangle \tilde{\lambda}_2 + \langle 43 \rangle \tilde{\lambda}_3 = 0$, which is the condition mandated by the first delta function above ($\langle 42 \rangle = -\langle 24 \rangle$). Contracting with λ_2 implies $\langle 21 \rangle \tilde{\lambda}_1 + \langle 23 \rangle \tilde{\lambda}_3 + \langle 24 \rangle \tilde{\lambda}_4 = 0$, which is the condition mandated by the second delta function. We therefore arrive at the result

$$M(1^+, 2^-, 3^+, 4^-) = \frac{\langle 24 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle} \, \delta^4\left(\sum_{i=1}^4 p_i\right) .$$

Recall that we wanted to show that the prefactor $A \equiv A(1^+, 2^-, 3^+, 4^-)$ of the delta function is equal to

$$\frac{[13]^4}{[12][23][34][41]} .$$

We will now show that the two forms are indeed equivalent. We will make repeated use of momentum conservation written in the form

$$|1\rangle[1| + |2\rangle[2| + |3\rangle[3| + |4\rangle[4| = 0 . \qquad (*)$$

Multiplying by $\langle 2|...|3]$ implies $\langle 24 \rangle = \langle 12 \rangle \frac{[13]}{[43]}$. We have

$$A = \frac{\langle 24 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle} = \frac{\langle 12 \rangle^4 [13]^4}{[43]^4 \langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle} .$$

Multiply $(*)$ by $\langle 2|...|4]$ to get $\langle 23 \rangle = \langle 12 \rangle \frac{[14]}{[34]}$. Therefore

$$A = \frac{\langle 12 \rangle^4 [13]^4 [34]}{[43]^4 \langle 12 \rangle^2 \langle 34 \rangle \langle 41 \rangle [14]} = \frac{\langle 12 \rangle^2 [13]^4}{[34]^3 \langle 34 \rangle \langle 41 \rangle [14]} .$$

Multiply $(*)$ by $\langle 4|...|2]$ to get $\langle 41 \rangle = \langle 34 \rangle \frac{[32]}{[12]}$ and therefore

$$A = \frac{\langle 12 \rangle^2 [13]^4 [12]}{[34]^3 \langle 34 \rangle^2 [14][32]} = \left(\frac{\langle 12 \rangle [12]}{\langle 34 \rangle [34]}\right)^2 \frac{[13]^4}{[12][34][14][32]} .$$

Since $(p_1 + p_2)^2 = (p_3 + p_4)^2$, the term in parentheses equals 1. Since $[14][32] = (-[41])(-[23]) = [41][23]$, we have

$$A = \frac{[13]^4}{[12][23][34][41]}$$

which is the desired result. Note that had we chosen to integrate over the delta functions in a different order we could have arrived at this form directly. It is however worth showing the relationship between the forms with angle and square brackets.

8. Show that $SL(4,R)$ is locally isomorphic to the conformal group. [Hint: Identify the $15 = 4^2 - 1$ generators of the conformal group (3 rotations J^i, 3 boosts K^i, 1 dilation D, 4 translations P^μ, and 4 conformal transformations K^μ) with the 15 traceless real 4 by 4 matrices.]

Solution:

Our strategy will be to write down all possible traceless real 4-by-4 matrices and identify them with the conformal generators based on their commutation relations. The 15 generators $M^{\mu\nu}$, P^μ, D, K^μ of the conformal group[47] satisfy the commutation relations:

$$[D, K^\mu] = -iK^\mu \ , \quad [D, P^\mu] = +iP^\mu \ , \quad [K^\mu, P^\nu] = +2i(\eta^{\mu\nu}D + M^{\mu\nu})$$
$$[K^\mu, M^{\nu\rho}] = -i(\eta^{\mu\nu}K^\rho - \eta^{\mu\rho}K^\nu) \ , \quad [P^\mu, M^{\nu\rho}] = -i(\eta^{\mu\nu}P^\rho - \eta^{\mu\rho}P^\nu)$$
$$[M^{\mu\nu}, M^{\rho\sigma}] = -i(\eta^{\nu\rho}M^{\mu\sigma} + \eta^{\mu\sigma}M^{\nu\rho} - \eta^{\mu\rho}M^{\nu\sigma} - \eta^{\nu\sigma}M^{\mu\rho})$$

with all others zero. The generators of rotations are $J^i = \frac{1}{2}\varepsilon^{ijk}M_{jk}$, and the generators of boosts are $B^i = M^{0i}$. (We will use B^i to denote the boosts so as not to confuse them with the spatial components of the special conformal generators.)

For two real spinors λ_α and $\mu^{\dot\alpha}$, we have learned in the text that conformal transformations act naturally on the 4-dimensional column vector

$$Z = \begin{pmatrix} \lambda_\alpha \\ \mu^{\dot\alpha} \end{pmatrix}$$

which transforms under the 4-dimensional representation of $SL(4,\mathbb{R})$. The lie algebra of $SL(4,\mathbb{R})$ consists of the 15 real traceless 4-by-4 matrices, which we now attempt to construct.

From learning about spinor representations of the Lorentz group back in Chapter II, we know how the rotations J^i and the boosts B^i act on the vector Z:

$$J^i = \begin{pmatrix} \frac{1}{2}\sigma^i & 0 \\ 0 & \frac{1}{2}\sigma^i \end{pmatrix} \ , \quad B^i = \begin{pmatrix} -i\frac{1}{2}\sigma^i & 0 \\ 0 & +i\frac{1}{2}\sigma^i \end{pmatrix} \ .$$

Here we are working in the signature $\eta = (-,+,+,+)$ of $SO(3,1)$. These generators can be repackaged into the relativistic notation

$$M^{\mu\nu} = \begin{pmatrix} \frac{1}{2}(\sigma^{\mu\nu})_\alpha{}^\beta & 0 \\ 0 & -\frac{1}{2}(\bar\sigma^{\mu\nu})^{\dot\alpha}{}_{\dot\beta} \end{pmatrix}$$

---

[47]Acting on scalar fields, these generators take the form:

$$M_{\mu\nu} = i(x_\mu\partial_\nu - x_\nu\partial_\mu)$$
$$P_\mu = -i\partial_\mu \ , \quad D = ix^\mu\partial_\mu$$
$$K_\mu = -i(x^2\partial_\mu - 2x_\mu x^\nu\partial_\nu) \ .$$

In addition to these, spinors and vectors also transform with the appropriate matrices. These matrices are what we are trying to identify.

where $\sigma^{\mu\nu} \equiv +\frac{1}{2}i(\sigma^\mu\bar\sigma^\nu - \sigma^\nu\bar\sigma^\mu)$ and $\bar\sigma^{\mu\nu} \equiv -\frac{1}{2}i(\bar\sigma^\mu\sigma^\nu - \bar\sigma^\nu\sigma^\mu)$, and we have defined the 4-vector

$$(\sigma^\mu)_{\alpha\dot\alpha} \equiv (I, \sigma^1, \sigma^2, \sigma^3) \ , \quad (\bar\sigma^\mu)^{\dot\alpha\alpha} \equiv \varepsilon^{\alpha\beta}\varepsilon^{\dot\alpha\dot\beta}\sigma^\mu_{\beta\dot\beta} = (I, -\sigma^1, -\sigma^2, -\sigma^3) \ .$$

These are defined so that

$$x_{\alpha\dot\alpha} \equiv \sigma^\mu_{\alpha\dot\alpha}x_\mu = \begin{pmatrix} x_0 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{pmatrix}$$

satisfies

$$\det x = x_0^2 - x_1^2 - x_2^2 - x_3^2 = -\eta^{\mu\nu}x_\mu x_\nu$$

with the $SO(3,1)$ metric $\eta^{\mu\nu} = \text{diag}(-1, +1, +1, +1)$. Note that the components of the matrices $\sigma^{\mu\nu}$ and $\bar\sigma^{\mu\nu}$ are numerically equal to:

$$\sigma^{ij} = +\varepsilon^{ijk}\sigma^k \ , \quad \sigma^{0i} = -i\sigma^i$$
$$\bar\sigma^{ij} = -\varepsilon^{ijk}\sigma^k \ , \quad \bar\sigma^{0i} = -i\sigma^i$$

with index placements as indicated by the form of $M^{\mu\nu}$ given above.

From looking at the index structure of these tensors as well as that of the spinors in $Z = (\lambda, \mu)$, the only other traceless 4-by-4 matrices we can write down are of the forms

$$X^\mu_\pm \equiv \begin{pmatrix} 0 & \frac{1}{2}\sigma^\mu_{\alpha\dot\beta} \\ \pm\frac{1}{2}\bar\sigma^{\mu\dot\alpha\beta} & 0 \end{pmatrix} \quad \text{and} \quad Y \equiv \begin{pmatrix} \frac{1}{2}\delta_\alpha{}^\beta & 0 \\ 0 & -\frac{1}{2}\delta^{\dot\alpha}{}_{\dot\beta} \end{pmatrix}$$

up to an overall constant for each. By direct computation, we find $[Y, X^\mu_\pm] = X^\mu_\mp$, and therefore:

$$[Y, X^\mu_\pm + X^\mu_\mp] = +(X^\mu_\pm + X^\mu_\mp)$$
$$[Y, X^\mu_\pm - X^\mu_\mp] = -(X^\mu_\pm - X^\mu_\mp)$$

We also find

$$[X^0_\pm, M^{ij}] = 0 \ , \quad [X^k_\pm, M^{ij}] = -i(\delta^{ik}X^j_\pm - \delta^{kj}X^i_\pm)$$
$$[X^i_\pm, M^{0j}] = i\delta^{ij}X^0_\pm \ , \quad [X^0_\pm, M^{0i}] = iX^i_\pm \ .$$

Finally, we also have

$$[X^\mu_+ + X^\mu_-, X^\nu_+ - X^\nu_-] = -2(\eta^{\mu\nu}Y + iM^{\mu\nu}) \ .$$

If we identify $D = iY$, $P^\mu = -i(X^\mu_+ + X^\mu_-)$ and $K^\mu = -i(X^\mu_+ - X^\mu_-)$, then we get the correct commutation relations for the conformal group. This completes the problem.

IX.2 N.3 Subterranean Connections in Gauge Theories

1. Show that the structure of Lie algebra (21) emerges naturally.

$$f_{abe}f_{cde} + f_{ace}f_{bde} + f_{ade}f_{bce} = 0 \qquad (21)$$

Solution:

First note that equation (21) on p. 509 has a sign error on the second term. The Jacobi identity should read

$$f_{abe}f_{cde} + f_{cae}f_{bde} + f_{ade}f_{bce} = 0 \ .$$

Recall that the recursion scheme is to separate the diagram into sets L and R, with $r \ \epsilon \ L$ and $s \ \epsilon \ R$. If $(r, s) = (1, 4)$, we get one separation with a pole in $s = (p_1 + p_2)^2 = (p_1 + p_3)^2$, and one with a pole in $t = (p_1 + p_3)^2 = (p_2 + p_4)^2$:

$$-\mathcal{M}(1_a, 2_b, 3_c, 4_d) = f_{abe}f_{ecd}A(1, 2, 3, 4) + f_{ace}f_{ebd}A(1, 3, 2, 4)$$

If $(r, s) = (1, 3)$, then we get one pole in s and one pole in $u = (p_1 + p_4)^2 = (p_2 + p_3)^2$:

$$-\mathcal{M}(1_a, 2_b, 3_c, 4_d) = f_{abe}f_{edc}A(1, 2, 4, 3) + f_{ade}f_{ebc}A(1, 4, 2, 3)$$

If $(r, s) = (1, 2)$, then we get one pole in t and one pole in u:

$$-\mathcal{M}(1_a, 2_b, 3_c, 4_d) = f_{ace}f_{edb}A(1, 3, 4, 2) + f_{ade}f_{ecb}A(1, 4, 3, 2)$$

To make sense of these, we should relate the various color-stripped amplitudes to each other. Since the amplitudes with all $+$, all $-$, or $3 +$ or $3 -$ are zero, we can without loss of generality consider the amplitude with two $+$ helicities and two $-$ helicities. From equation (19) on p. 493 and the surrounding discussion, we know how to compute the amplitude for any ordering of the $+$ and $-$ helicities, so without loss of generality we choose $(1^-, 2^-, 3^+, 4^+)$. Thus consider the amplitude

$$A(1^-, 2^-, 3^+, 4^+) = \frac{\langle 12 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle}$$

from equation (17) on p. 492. The goal is to relate all of the other As to this one. We will repeatedly use momentum conservation $\sum_{i=1}^{4} p_i = 0$ in the form

$$|1\rangle[1| + |2\rangle[2| + |3\rangle[3| + |4\rangle[4| = 0 \ . \qquad (*)$$

Consider the amplitude $A(1^-, 3^+, 2^-, 4^+)$. From relabeling equation (19) on p. 493, we know

$$A(1^-, 3^+, 2^-, 4^+) = \frac{\langle 12 \rangle^4}{\langle 13 \rangle \langle 32 \rangle \langle 24 \rangle \langle 41 \rangle} = \left(\frac{\langle 12 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle} \right) \left(\frac{-\langle 12 \rangle \langle 34 \rangle}{\langle 13 \rangle \langle 24 \rangle} \right) \ .$$

Multiply the momentum conservation equation $(*)$ by $\langle 1|$ on the left and by $|4]$ on the right to get

$$\langle 12 \rangle[24] + \langle 13 \rangle[34] = 0 \implies \frac{\langle 12 \rangle}{\langle 13 \rangle} = -\frac{[34]}{[24]} \ .$$

Thus
$$\frac{-\langle 12\rangle\langle 34\rangle}{\langle 13\rangle\langle 24\rangle} = \frac{+[34]\langle 34\rangle}{[24]\langle 24\rangle} = \frac{s}{t} \ .$$

Therefore, we have
$$A(1^-,3^+,2^-,4^+) = \frac{s}{t}\, A(1^-,2^-,3^+,4^+)\ .$$

Next, consider
$$A(1^-,2^-,4^+,3^+) = \frac{\langle 12\rangle^4}{\langle 12\rangle\langle 24\rangle\langle 43\rangle\langle 31\rangle} = \left(\frac{\langle 12\rangle^4}{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle}\right)\left(\frac{-\langle 23\rangle\langle 41\rangle}{\langle 24\rangle\langle 31\rangle}\right)$$

Multiply $(*)$ by $\langle 2|$ and $|1]$ to get $\langle 23\rangle[31] + \langle 24\rangle[41] = 0 \implies \langle 23\rangle/\langle 24\rangle = -[41]/[31]$ and thus
$$\frac{-\langle 23\rangle\langle 41\rangle}{\langle 24\rangle\langle 31\rangle} = \frac{[41]\langle 41\rangle}{[31]\langle 31\rangle} = \frac{u}{t} \ .$$

We have
$$A(1^-,2^-,4^+,3^+) = \frac{u}{t}\, A(1^-,2^-,3^+,4^+)\ .$$

Next, consider
$$A(1^-,4^+,2^-,3^+) = \frac{\langle 12\rangle^4}{\langle 14\rangle\langle 42\rangle\langle 23\rangle\langle 31\rangle} = \left(\frac{\langle 12\rangle^4}{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle}\right)\left(\frac{-\langle 12\rangle\langle 34\rangle}{\langle 42\rangle\langle 31\rangle}\right)\ .$$

Multiply $(*)$ by $\langle 2|$ and $[3|$ to get $\langle 21\rangle[13] + \langle 24\rangle[43] = 0 \implies \langle 12\rangle/\langle 42\rangle = -[34]/[31]$ and thus
$$\frac{-\langle 12\rangle\langle 34\rangle}{\langle 42\rangle\langle 31\rangle} = \frac{+[34]\langle 34\rangle}{[31]\langle 31\rangle} = \frac{s}{t} \ .$$

We have
$$A(1^-,4^+,2^-,3^+) = \frac{s}{t}\, A(1^-,2^-,3^+,4^+)\ .$$

Two more to go. The next one is
$$A(1^-,3^+,4^+,2^-) = \frac{\langle 12\rangle^4}{\langle 13\rangle\langle 34\rangle\langle 42\rangle\langle 21\rangle} = \left(\frac{\langle 12\rangle^4}{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle}\right)\left(\frac{-\langle 23\rangle\langle 41\rangle}{\langle 13\rangle\langle 42\rangle}\right)\ .$$

Multiply $(*)$ by $\langle 3|$ and $|4]$ to get $\langle 31\rangle[14] + \langle 32\rangle[24] = 0 \implies \langle 23\rangle/\langle 13\rangle = -[41]/[42]$ and thus
$$\frac{-\langle 23\rangle\langle 41\rangle}{\langle 13\rangle\langle 42\rangle} = \frac{+[41]\langle 41\rangle}{[42]\langle 42\rangle} = \frac{u}{t} \ .$$

We have
$$A(1^-,3^+,4^+,2^-) = \frac{u}{t}\, A(1^-,2^-,3^+,4^+)\ .$$

Fortunately, the last one is easy:
$$A(1^-,4^+,3^+,2^-) = \frac{\langle 12\rangle^4}{\langle 14\rangle\langle 43\rangle\langle 32\rangle\langle 21\rangle} = \frac{\langle 12\rangle^4}{\langle 12\rangle\langle 23\rangle\langle 34\rangle\langle 41\rangle} = A(1^-,2^-,3^+,4^+)\ .$$

We therefore arrive at the amplitude

$$-\mathcal{M}(1_a^-, 2_b^-, 3_c^+, 4_d^+) = \left(f_{abe} f_{ecd} + \frac{s}{t} f_{ace} f_{ebd} \right) A(1^-, 2^-, 3^+, 4^+)$$

$$= \left(\frac{u}{t} f_{abe} f_{edc} + \frac{s}{t} f_{ade} f_{ebc} \right) A(1^-, 2^-, 3^+, 4^+)$$

$$= \left(\frac{u}{t} f_{ace} f_{edb} + f_{ade} f_{ecb} \right) A(1^-, 2^-, 3^+, 4^+) .$$

The equalities reflect the physical necessity for the amplitude to be the same regardless of which (r, s) we choose for the recursion formula. Subtracting and multiplying by t gives the consistency conditions

$$t f_{abe} f_{ecd} + s(f_{ace} f_{ebd} - f_{ade} f_{ebc}) - u f_{abe} f_{edc} = 0$$
$$t(f_{abe} f_{ecd} - f_{ade} f_{ecb}) + s f_{ace} f_{ebd} - u f_{ace} f_{edb} = 0$$

Subtract (2) from (1) to get

$$t f_{ade} f_{ecb} - s f_{ade} f_{ebc} + u(f_{ace} f_{edb} - f_{abe} f_{edc}) = 0 .$$

Since $f_{ecb} = -f_{ebc}$ and since $s + t + u = 0 \implies t + s = -u$, we get

$$u(f_{ade} f_{ebc} + f_{ace} f_{edb} - f_{abe} f_{edc}) = 0 .$$

$u \neq 0$ and $f_{edc} = -f_{ecd}$ implies

$$f_{ade} f_{ebc} + f_{ace} f_{edb} + f_{abe} f_{ecd} = 0 .$$

This is the Jacobi identity.

2. In appendix 1 we recursed by complexifying the momenta of two external lines with helicity $+$ and $-$. In the derivation of the recursion relation (3) we could have picked any two external lines to complexify. Determine the amplitude calculated directly in chapter N.2, namely $A(1^-, 2^-, 3^+, 4^+)$, by complexifying lines 1 and 2. This is an example of the self-consistency argument sketched in the text.

Particle physics experimentalists are fond of saying that yesterday's spectacular discovery is today's calibration and tomorrow's annoying background. The canonical example is the Nobel-winning discovery of the CP-violating decay of the K_L meson into two pions. In theoretical physics, yesterday's discovery is today's homework exercise and tomorrow's trivium.

$$\mathcal{M}(0) = -\sum_{L,h} \frac{\mathcal{R}_L}{z_L} = -\sum_{L,h} \frac{\mathcal{M}_L(z_L) \mathcal{M}_R(z_L)}{P_L(0)^2} \tag{3}$$

Solution:

We are interested in the color-ordered amplitude $A(1^-, 2^-, 3^+, 4^+)$ with legs 1 and 2 complexified. There is only one way to split up the amplitude into 3-point vertices for which legs 1 and 2 do not belong to the same partition while maintaining the color ordering:
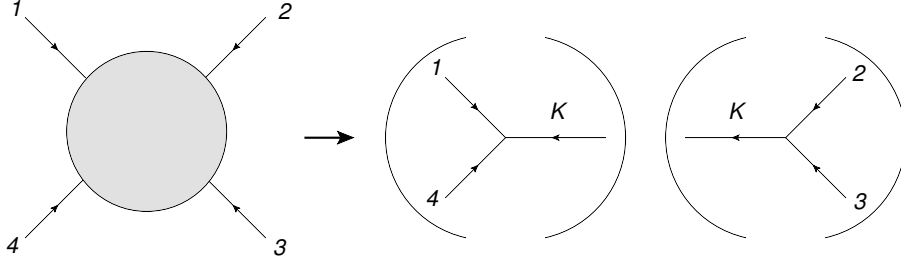
The color-ordered amplitude is:

$$A(1^-, 2^-, 3^+, 4^+) = - \sum_{h=\pm} A(\hat{1}^-, \hat{K}^h, 4^+) \frac{1}{K^2} A(-\hat{K}^{-h}, \hat{2}^-, 3^+)$$

$$= -\frac{1}{K^2} \left[A(\hat{K}^+, 4^+, \hat{1}^-) A(-\hat{K}^-, \hat{2}^-, 3^+) + A(\hat{1}^-, \hat{K}^-, 4^+) A(3^+, -\hat{K}^+, \hat{2}^-) \right]$$

In the second line we have cyclically permuted the arguments of the 3-point amplitudes to put them into the canonical forms:

$$A(1^+, 2^+, 3^+) = \frac{[12]^3}{[23][31]} , \quad A(1^-, 2^-, 3^+) = \frac{\langle 12 \rangle^3}{\langle 23 \rangle \langle 31 \rangle} .$$

The internal momentum is $K = -(p_1 + p_4) = +(p_2 + p_3)$, which in spinor notation reads

$$|K\rangle[K| = -|1\rangle[1| - |4\rangle[4| = |2\rangle[2| + |3\rangle[3| . \qquad (A)$$

When the internal momentum $K \equiv |K\rangle[K|$ flows out of the vertex, we can effect the replacement $K \to -K$ by replacing its spinors as $|K\rangle \to i|K\rangle$, $|K] \to i|K]$. We have:

$$A(1^-, 2^-, 3^+, 4^+) = - \frac{1}{2\langle 23 \rangle [23]} \left[\left(\frac{[\hat{K}4]^3}{[4\hat{1}][\hat{1}\hat{K}]} \right) \left(\frac{(i\langle \hat{K}\hat{2} \rangle)^3}{\langle \hat{2}3 \rangle (i\langle 3\hat{K} \rangle)} \right) + \left(\frac{\langle \hat{1}\hat{K} \rangle^3}{\langle \hat{K}4 \rangle \langle 4\hat{1} \rangle} \right) \left(\frac{(i[3\hat{K}])^3}{(i[\hat{K}\hat{2}])[\hat{2}3]} \right) \right]$$

$$= + \frac{1}{2\langle 23 \rangle [23]} \left[\frac{[\hat{K}4]^3}{[4\hat{1}][\hat{1}\hat{K}]} \frac{\langle \hat{K}\hat{2} \rangle^3}{\langle \hat{2}3 \rangle \langle 3\hat{K} \rangle} + \frac{\langle \hat{1}\hat{K} \rangle^3}{\langle \hat{K}4 \rangle \langle 4\hat{1} \rangle} \frac{[3\hat{K}]^3}{[\hat{K}\hat{2}][\hat{2}3]} \right] .$$

Here we choose $(r, s) = (1, 2)$ and thereby complexify the momenta as:

$$|\hat{1}] = |1] + z_L |2] , \quad |\hat{2}\rangle = |2\rangle - z_L |1\rangle$$

with $|1\rangle$ and $|2]$ unchanged. This effects the shift $\hat{p}_1 = p_1 + z_L q , \quad \hat{p}_2 = p_2 - z_L q$ with $q = |1\rangle[2|$.

Here we have removed the hats from $|1\rangle$ and $|2]$ since they are not complexified. Now use the hatted version of equation (A):

$$\langle \hat{2}|(\hat{A})|4] \implies \langle \hat{2}\hat{K} \rangle [\hat{K}4] = -\langle 21 \rangle [4\hat{1}]$$
$$\langle 3|(\hat{A})|\hat{1}] \implies \langle 3\hat{K} \rangle [\hat{K}\hat{1}] = -\langle 34 \rangle [4\hat{1}]$$

314

Therefore:

$$\frac{[\hat{K}4]^3 \langle \hat{K}\hat{2}\rangle^3}{[4\hat{1}][\hat{1}\hat{K}]\langle\hat{2}3\rangle\langle 3\hat{K}\rangle} = \frac{(\langle 21\rangle)[4\hat{1}])^3}{[4\hat{1}]\langle\hat{2}3\rangle(\langle 34\rangle[4\hat{1}])} = \frac{\langle 21\rangle^3[4\hat{1}]}{\langle\hat{2}3\rangle\langle 34\rangle} \ .$$

Conservation of momentum implies

$$|1\rangle[1| + |2\rangle[2| + |3\rangle[3| + |4\rangle[4| = 0 \ .$$

Multiply the hatted version of this by $\langle\hat{2}|...|4]$ to obtain

$$\frac{[\hat{1}4]}{\langle\hat{2}3\rangle} = -\frac{[34]}{\langle 21\rangle}$$

and therefore

$$\frac{\langle 21\rangle^3[4\hat{1}]}{\langle\hat{2}3\rangle\langle 34\rangle} = \frac{\langle 21\rangle^2[34]}{\langle 34\rangle} = \frac{\langle 21\rangle^3}{\langle 23\rangle\langle 34\rangle}\left(\frac{[34]}{\langle 21\rangle[23]}\right) \ .$$

Multiplying momentum conservation (unhatted) by $\langle 1|...|3]$ implies that the term in parentheses equals $-\langle 41\rangle$, so all together we have[48]

$$\frac{[\hat{K}4]^3\langle\hat{K}\hat{2}\rangle^3}{[4\hat{1}][\hat{1}\hat{K}]\langle\hat{2}3\rangle\langle 3\hat{K}\rangle} = \frac{\langle 12\rangle^3}{\langle 23\rangle\langle 34\rangle\langle 41\rangle} \ .$$

This is half of the correct answer.

Now let us simplify the second term. Go back to equation (A) and obtain:

$$\langle 1|(\hat{A})|3] \implies \langle 1\hat{K}\rangle[\hat{K}3] = +\langle 12\rangle[23]$$
$$\langle 4|(\hat{A})|2] \implies \langle 4\hat{K}\rangle[\hat{K}2] = -\langle 41\rangle[12] \ .$$

Therefore:

$$\frac{1}{\langle 23\rangle[23]}\frac{\langle\hat{1}\hat{K}\rangle^3}{\langle\hat{K}4\rangle\langle 4\hat{1}\rangle}\frac{[3\hat{K}]^3}{[\hat{K}\hat{2}][\hat{2}3]} = \frac{1}{\langle 23\rangle[23]}\left(\frac{-\langle 12\rangle^3[23]^3}{\langle 41\rangle[23]}\right)\left(\frac{1}{+\langle 41\rangle[12]}\right)$$
$$= \frac{-\langle 12\rangle^3}{\langle 23\rangle\langle 41\rangle}\frac{[23]}{\langle 41\rangle[12]} \ .$$

Multiply momentum conservation (unhatted) by $\langle 4|...|2]$ to get

$$\langle 41\rangle[12] + \langle 43\rangle[32] = 0 \implies \frac{\langle 41\rangle[12]}{[23]} = -\langle 34\rangle$$

so that the above term equals

$$\frac{\langle 12\rangle^3}{\langle 23\rangle\langle 34\rangle\langle 41\rangle} \ .$$

---

[48]We thank Eric Dzienkowski for catching an error here.

This is the second half of the correct answer. Adding the two terms cancels the overall factor of 2, and we arrive at the result

$$A(1^-, 2^-, 3^+, 4^+) = \frac{\langle 12 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle} \ .$$

3. Using the explicit forms given for $A(1^-, 2^-, 3^+, 4^+)$ and $A(1^-, 2^+, 3^-, 4^+)$ in the preceding chapter, check the estimated large z behavior in (13-15).

$$\mathcal{M}^{-+}(z) \to \frac{1}{z} \qquad\qquad (13)$$

$$\mathcal{M}^{--}(z) \to \frac{1}{z} \qquad\qquad (14)$$

$$\mathcal{M}^{+-}(z) \to z^3 \qquad\qquad (15)$$

Solution:

The amplitudes are given on pages 492-493:

$$A(1^-, 2^-, 3^+, 4^+) = \frac{\langle 12 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle} \ , \quad A(1^-, 2^+, 3^-, 4^+) = \frac{\langle 13 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle}$$

and the physical amplitude specified by helicity labels for particles r and s is $\mathcal{M}^{h_r h_s}(z) \equiv [\epsilon_r^{h_r}(z)]_\mu \mathcal{M}^{\mu\nu}(z) [\epsilon_s^{h_s}(z)]_\nu$. The spinorial version of the deformations $p_r(z) = p_r + zq$ and $p_s(z) = p_s - zq$ are given on p. 501, with $p_i = \lambda_i \tilde{\lambda}_i$:

$$\lambda_r(z) = \lambda_r \ , \quad \tilde{\lambda}_r(z) = \tilde{\lambda}_r + z \tilde{\lambda}_s$$
$$\lambda_s(z) = \lambda_s - z \lambda_r \ , \quad \tilde{\lambda}_s(z) = \tilde{\lambda}_s$$

and $q = \lambda_r \tilde{\lambda}_s$. Choose $(r, s) = (1, 2)$. Then the only z-dependence in the amplitude $A(1^-, 2^+, 3^-, 4^+)$ potentially comes from $\langle 12 \rangle$ and $\langle 23 \rangle$. But $\langle 12 \rangle(z) = \langle 12 \rangle - z \langle 11 \rangle = \langle 12 \rangle$, so really the only dependence on z comes from $\langle 23 \rangle(z) = \langle 23 \rangle - z \langle 13 \rangle \sim -z \langle 13 \rangle$. Therefore we find $\mathcal{M}^{-+}(z) \sim 1/z$.

The case for $A(1^-, 2^-, 3^+, 4^+)$ is identical, with the only z-dependence coming from a factor of $\langle 23$ in the denominator, leading immediately to $\mathcal{M}^{--}(z) \sim 1/z$.

Now choose $(r, s) = (2, 3)$. Then $\langle 13 \rangle(z) = \langle 13 \rangle - z \langle 12 \rangle \sim z$, leading to a factor of z^4 in the numerator. This is canceled partially by a factor $\langle 34 \rangle(z) = \langle 34 \rangle - z \langle 14 \rangle \sim z$ in the denominator, leading to $\mathcal{M}^{+-}(z) \sim z^3$.

4. Worry about the sloppy handling of factors of 2 in appendix 1. [Hint: The final result is correct because the polarization vectors in (5-6) are normalized to $|\varepsilon|^2 = 2$ for convenience.]

Solution:

Let us be pedantic about distinguishing among p^μ, $p_{\alpha\dot\alpha}$ and $p^{\dot\alpha\alpha}$. Given the Lorentz 4-vector p^μ, define $\not{p}_{\alpha\dot\alpha} \equiv p_\mu \sigma^\mu_{\alpha\dot\alpha}$ and $\not{\bar{p}}^{\dot\alpha\alpha} \equiv p_\mu \bar\sigma^{\mu\dot\alpha\alpha}$, where $\bar\sigma^{\mu\dot\alpha\alpha} \equiv \varepsilon^{\alpha\beta}\varepsilon^{\dot\alpha\dot\beta}\sigma^\mu_{\beta\dot\beta}$. Then for lightlike momenta $p^2 \equiv p_\mu p^\mu = 0$, we define $\not{p}_{\alpha\dot\alpha} \equiv \lambda_\alpha \tilde\lambda_{\dot\alpha}$. Then $\lambda^\alpha \tilde\lambda^{\dot\alpha} = \varepsilon^{\alpha\beta}\varepsilon^{\dot\alpha\dot\beta}\lambda_\beta\tilde\lambda_{\dot\beta} = \varepsilon^{\alpha\beta}\varepsilon^{\dot\alpha\dot\beta}\not{p}_{\beta\dot\beta} = \not{\bar{p}}^{\dot\alpha\alpha}$. Therefore:

$$\langle\lambda\lambda'\rangle[\lambda\lambda'] = \lambda_\alpha\lambda'^\alpha\tilde\lambda_{\dot\alpha}\tilde\lambda'^{\dot\alpha} = \not{p}_{\alpha\dot\alpha}\,\not{\bar{p}}^{\dot\alpha\alpha} = p_\mu p_\nu \mathrm{tr}(\sigma^\mu\bar\sigma^\nu) = 2\,p_\mu p^\mu \ .$$

So the top of page 511 should read $P_L(0)^2 \equiv P_L(0)_\mu P_L(0)^\mu = 2\,p_{1\mu}p_2^\mu = \langle 12\rangle[12]$. Thus the result (25) is correct.

X   Appendix E: Dotted and Undotted Indices

1. Show that $\eta\sigma^{\mu\nu}\psi = -\psi\sigma^{\mu\nu}\eta$ and $\bar{\chi}\bar{\sigma}^\mu\psi = -\psi\sigma^\mu\bar{\chi}$.

Solution: First consider the second object. We have:

$$
\begin{aligned}
\bar{\chi}\bar{\sigma}^\mu\psi &= \bar{\chi}_{\dot{\alpha}}\bar{\sigma}^{\mu\dot{\alpha}\alpha}\psi_\alpha \\
&= -\psi_\alpha\bar{\sigma}^{\mu\dot{\alpha}\alpha}\bar{\chi}_{\dot{\alpha}} \\
&= -\varepsilon_{\alpha\beta}\psi^\beta\bar{\sigma}^{\mu\dot{\alpha}\alpha}\varepsilon_{\dot{\alpha}\dot{\beta}}\bar{\chi}^{\dot{\beta}} \\
&= -\psi^\beta(-1)^2\left(\varepsilon_{\beta\alpha}\varepsilon_{\dot{\beta}\dot{\alpha}}\bar{\sigma}^{\mu\dot{\alpha}\alpha}\right)\bar{\chi}^{\dot{\beta}} \\
&= -\psi^\beta\sigma^\mu_{\beta\dot{\beta}}\bar{\chi}^{\dot{\beta}} \\
&= -\psi\sigma^\mu\bar{\chi} \quad\checkmark
\end{aligned}
$$

For the first object, recall that $\sigma^{\mu\nu} = \frac{1}{4}(\sigma^\mu\bar{\sigma}^\nu - \sigma^\nu\bar{\sigma}^\mu)$. Therefore:

$$
\begin{aligned}
4\,\eta\sigma^{\mu\nu}\psi &= \eta\sigma^\mu\bar{\sigma}^\nu\psi - (\mu\leftrightarrow\nu) = \eta^\alpha\sigma^\mu_{\alpha\dot{\beta}}\bar{\sigma}^{\nu\dot{\beta}\beta}\psi_\beta - (\mu\leftrightarrow\nu) \\
&= -\psi_\beta\sigma^\mu_{\alpha\dot{\beta}}\bar{\sigma}^{\nu\dot{\beta}\beta}\eta^\alpha - (\mu\leftrightarrow\nu) \\
&= -\psi_\beta[(\varepsilon_{\alpha\gamma}\varepsilon_{\dot{\beta}\dot{\gamma}}\bar{\sigma}^{\mu\dot{\gamma}\gamma})(\varepsilon^{\dot{\beta}\dot{\delta}}\varepsilon^{\beta\delta}\sigma^\nu_{\delta\dot{\delta}})]\eta^\alpha - (\mu\leftrightarrow\nu) \\
&= -\psi_\beta[\varepsilon_{\alpha\gamma}\varepsilon^{\beta\delta}(-\delta^{\dot{\delta}}_{\dot{\gamma}})\bar{\sigma}^{\mu\dot{\gamma}\gamma}\sigma^\nu_{\delta\dot{\delta}}]\eta^\alpha - (\mu\leftrightarrow\nu) \qquad [\text{ note: } \varepsilon_{\dot{\beta}\dot{\gamma}}\varepsilon^{\dot{\beta}\dot{\delta}} = -\varepsilon_{\dot{\gamma}\dot{\beta}}\varepsilon^{\dot{\beta}\dot{\delta}} = -\delta^{\dot{\delta}}_{\dot{\gamma}}] \\
&= -\psi_\beta(-\varepsilon_{\alpha\gamma}\varepsilon^{\beta\delta}\bar{\sigma}^{\mu\dot{\gamma}\gamma}\sigma^\nu_{\delta\dot{\delta}})\eta^\alpha - (\mu\leftrightarrow\nu) \\
&= +(\varepsilon^{\beta\delta}\psi_\beta)(\bar{\sigma}^{\mu\dot{\gamma}\gamma}\sigma^\nu_{\delta\dot{\delta}})(\varepsilon_{\alpha\gamma}\eta^\alpha) - (\mu\leftrightarrow\nu) \\
&= (-\psi^\delta)(\bar{\sigma}^{\mu\dot{\gamma}\gamma}\sigma^\nu_{\delta\dot{\delta}})(-\eta_\gamma) - (\mu\leftrightarrow\nu) \\
&= +\psi\sigma^\nu\bar{\sigma}^\mu\eta - (\mu\leftrightarrow\nu) \qquad [\text{ put in matrix multiplication order }] \\
&= -[\psi\sigma^\mu\bar{\sigma}^\nu\eta - (\mu\leftrightarrow\nu)] \\
&= -4\,\psi\sigma^{\mu\nu}\eta \implies \eta\sigma^{\mu\nu}\psi = -\psi\sigma^{\mu\nu}\eta \quad\checkmark
\end{aligned}
$$

2. Show that $(\theta\varphi)(\bar{\chi}\bar{\xi}) = -\frac{1}{2}(\theta\sigma^\mu\bar{\xi})(\bar{\chi}\bar{\sigma}_\mu\varphi)$.

Solution:

First we need
$$
\sigma^\mu_{\alpha\dot{\alpha}}\bar{\sigma}^{\dot{\beta}\beta}_\mu = C\delta_\alpha{}^\beta\delta_{\dot{\alpha}}{}^{\dot{\beta}}
$$

which we know by $SU(2)\otimes SU(2)$ invariance. Choose $\alpha = \beta = \dot{\alpha} = \dot{\beta} = 1$ to get

$$
C = \sigma^\mu_{11}\bar{\sigma}^{11}_\mu = (\sigma^0)_{11}(\bar{\sigma}^0)^{11} - (\sigma^3)_{11}(\bar{\sigma}^3)^{11} = 1 - (-1) = 2
$$

Therefore

$$
\sigma^\mu_{\alpha\dot{\alpha}}\bar{\sigma}^{\dot{\beta}\beta}_\mu = 2\,\delta_\alpha{}^\beta\delta_{\dot{\alpha}}{}^{\dot{\beta}}\ .
$$

Now compute:

$$(\theta\sigma^\mu\bar\xi)(\bar\chi\bar\sigma_\mu\varphi) = \theta^\alpha\sigma^\mu_{\alpha\dot\alpha}\bar\xi^{\dot\alpha}\bar\chi_{\dot\beta}\bar\sigma_\mu^{\dot\beta\beta}\varphi_\beta$$
$$= 2\,\theta^\alpha\bar\xi^{\dot\alpha}\delta_\alpha{}^\beta\delta_{\dot\alpha}{}^{\dot\beta}\bar\chi_{\dot\beta}\varphi_\beta$$
$$= 2\,\theta^\alpha\bar\xi^{\dot\alpha}\bar\chi_{\dot\alpha}\varphi_\alpha$$
$$= 2(-1)\,\theta^\alpha\bar\chi_{\dot\alpha}\bar\xi^{\dot\alpha}\varphi_\alpha$$
$$= 2(-1)(-1)^2\theta^\alpha\varphi_\alpha\bar\chi_{\dot\alpha}\bar\xi^{\dot\alpha}$$
$$= -2\,(\theta\varphi)(\bar\chi\bar\xi)\ .$$

Dividing by 2 proves the statement.

3. Show that $\theta^\alpha\theta_\beta = \frac{1}{2}(\theta\theta)\delta^\alpha{}_\beta$. [Hint: simply evaluate the two sides for all possible cases.]

Solution:

The indices have to match up by $SU(2)$ invariance, and then we can fix the constant.

$$\theta^\alpha\theta_\beta = C(\theta\theta)\delta^\alpha{}_\beta$$

Contract both sides with $\delta^\beta{}_\alpha$ and note that $\delta^\alpha{}_\alpha = 2$ to get

$$(\theta\theta) = C(\theta\theta)\delta^\alpha{}_\alpha = 2C(\theta\theta) \implies C = \frac{1}{2}\ .$$

Alternatively, we can follow the hint and just check. $\alpha = 1, \beta = 2$ implies 0 for the right-hand side, and

$$\theta^1\theta_2 = \varepsilon^{12}\theta_2\theta_2 = 0$$

for the left-hand side, since $\theta_2^2 = 0$ since the square of any Grassmann number is zero. For $\alpha = \beta = 1$, we have $\frac{1}{2}(\theta\theta) = \frac{1}{2}(\theta^1\theta_1 + \theta^2\theta_2)$, and $\theta^2\theta_2 = \varepsilon^{21}\theta_1\varepsilon_{21}\theta^1 = -(\varepsilon^{21})^2\theta_1\theta^1 = -\theta_1\theta^1 = +\theta^1\theta_1$, so $\frac{1}{2}(\theta\theta) = \theta^1\theta_1$, which matches the left-hand side $\theta^\alpha\theta_\beta$ for $\alpha = \beta = 1$.

Addendum: More on Two-Component Spinors

Here we collect a few relations between two-component spinors and four-component spinors for use in calculating Feynman diagrams using two-component notation. For an extensive review, consult arXiv:0812.1594v5 [hep-ph].

Take the Dirac field for the electron for definiteness: $E(x) \equiv \begin{pmatrix} e_\alpha(x) \\ \bar e^{\dagger\dot\alpha}(x) \end{pmatrix}$. The mode expansions for the fields $e_\alpha(x)$ and $\bar e_\alpha(x)$ are:

$$e_\alpha(x) = \sum_s \int \frac{d^3p}{(2\pi)^3 2\omega_p}\left[\,b(\vec p, s)u_\alpha(\vec p, s)\,e^{-ip\cdot x} + d^\dagger(\vec p, s)v_\alpha(\vec p, s)\,e^{+ip\cdot x}\right]$$

$$\bar e_\alpha(x) = \sum_s \int \frac{d^3p}{(2\pi)^3 2\omega_p}\left[\,d(\vec p, s)u_\alpha(\vec p, s)\,e^{-ip\cdot x} + b^\dagger(\vec p, s)v_\alpha(\vec p, s)\,e^{+ip\cdot x}\right]\ .$$

319

The creation operators $b^\dagger$ and $d^\dagger$ create an electron e^- and a positron e^+ respectively.

The wave functions $u(\vec{p}, s)$ and $v(\vec{p}, s)$ are related to the 4-component wave functions $U(\vec{p}, s)$ and $V(\vec{p}, s)$ as

$$U(\vec{p}, s) = \begin{pmatrix} u_\alpha(\vec{p}, s) \\ v^{\dagger\dot\alpha}(\vec{p}, s) \end{pmatrix} , \quad V(\vec{p}, s) = \begin{pmatrix} v_\alpha(\vec{p}, s) \\ u^{\dagger\dot\alpha}(\vec{p}, s) \end{pmatrix}$$

The usual spin sum relations (see p. 110 of the text) $\sum_s U(\vec{p}, s)\bar{U}(\vec{p}, s) = \not{p} + m$ and $\sum_s V(\vec{p}, s)\bar{V}(\vec{p}, s) = \not{p} - m$, along with the gamma matrices in the Weyl basis (see appendix E)

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}$$

imply the spin sum relations

$$\sum_s \begin{pmatrix} u_\alpha v^\beta & u_\alpha u^\dagger_{\dot\beta} \\ v^{\dagger\dot\alpha} v^\beta & v^{\dagger\dot\alpha} u^\dagger_{\dot\beta} \end{pmatrix} = \begin{pmatrix} m\, \delta_\alpha{}^\beta & \not{p}_{\alpha\dot\beta} \\ \bar{\not{p}}^{\dot\alpha\beta} & m\, \delta^{\dot\alpha}{}_{\dot\beta} \end{pmatrix}$$

$$\sum_s \begin{pmatrix} v_\alpha u^\beta & v_\alpha v^\dagger_{\dot\beta} \\ u^{\dagger\dot\alpha} u^\beta & u^{\dagger\dot\alpha} v^\dagger_{\dot\beta} \end{pmatrix} = \begin{pmatrix} -m\, \delta_\alpha{}^\beta & \not{p}_{\alpha\dot\beta} \\ \bar{\not{p}}^{\dot\alpha\beta} & -m\, \delta^{\dot\alpha}{}_{\dot\beta} \end{pmatrix}$$

where we have defined $\not{p}_{\alpha\dot\beta} \equiv \sigma^\mu_{\alpha\dot\beta} p_\mu$ and $\bar{\not{p}}^{\dot\alpha\beta} \equiv \bar{\sigma}^{\mu\dot\alpha\beta} p_\mu$. Since $\mathrm{tr}(\sigma^\mu \bar{\sigma}^\nu) = 2\eta^{\mu\nu}$ [with $\eta = (+, -, -, -)$] we have $\not{p}_{\alpha\dot\beta} \bar{\not{p}}'^{\dot\beta\alpha} = 2p \cdot p'$.

An incoming electron state with momentum $\vec{p}$ and spin s is $|e^-(\vec{p}, s)\rangle = (2\pi)^3 2\omega_p b^\dagger(\vec{p}, s)|0\rangle$. This implies the following rules for the external state wave function for an incoming electron:

$$\langle 0|e(x)|e^-(\vec{p}, s)\rangle = u(\vec{p}, s)\, e^{-ip\cdot x} , \quad \langle 0|\bar{e}(x)|e^-(\vec{p}, s)\rangle = 0$$
$$\langle 0|e^\dagger(x)|e^-(\vec{p}, s)\rangle = 0 , \quad \langle 0|\bar{e}^\dagger(x)|e^-(\vec{p}, s)\rangle = v^\dagger(\vec{p}, s)\, e^{-ip\cdot x} .$$

An outgoing electron state with momentum $\vec{p}$ and spin s is $\langle e^-(\vec{p}, s)| = \langle 0|b(\vec{p}, s)(2\pi)^3 2\omega_p$. This implies the following rules for the external state wave function for an outgoing electron:

$$\langle e^-(\vec{p}, s)|e(x)|0\rangle = 0 , \quad \langle e^-(\vec{p}, s)|\bar{e}(x)|0\rangle = v(\vec{p}, s)\, e^{+ip\cdot x}$$
$$\langle e^-(\vec{p}, s)|e^\dagger(x)|0\rangle = u^\dagger(\vec{p}, s)\, e^{+ip\cdot x} , \quad \langle e^-(\vec{p}, s)|\bar{e}^\dagger(x)|0\rangle = 0 .$$

An incoming positron state with momentum $\vec{p}$ and spin s is $|e^+(\vec{p}, s)\rangle = (2\pi)^3 2\omega_p d^\dagger(\vec{p}, s)|0\rangle$. This implies the following rules for the external state wave function for an incoming positron:

$$\langle 0|e(x)|e^+(\vec{p}, s)\rangle = 0 , \quad \langle 0|\bar{e}(x)|e^+(\vec{p}, s)\rangle = u(\vec{p}, s)\, e^{-ip\cdot x}$$
$$\langle 0|e^\dagger(x)|e^+(\vec{p}, s)\rangle = v^\dagger(\vec{p}, s)\, e^{-ip\cdot x} , \quad \langle 0|\bar{e}^\dagger(x)|e^+(\vec{p}, s)\rangle = 0 .$$

An outgoing positron state with momentum $\vec{p}$ and spin s is $\langle e^+(\vec{p}, s)| = \langle 0|d(\vec{p}, s)(2\pi)^3 2\omega_p$. This implies the following rules for the external state wave function for an incoming positron:

$$\langle e^+(\vec{p}, s)|e(x)|0\rangle = v(\vec{p}, s)\, e^{+ip\cdot x} , \quad \langle e^+(\vec{p}, s)|\bar{e}(x)|0\rangle = 0$$
$$\langle e^+(\vec{p}, s)|e^\dagger(x)|0\rangle = 0 , \quad \langle e^+(\vec{p}, s)|\bar{e}^\dagger(x)|0\rangle = u^\dagger(\vec{p}, s)\, e^{+ip\cdot x} .$$